Anna van Buerenplein 1
2595 DA Den Haag
P.O. Box 96800
2509 JE The Hague
The Netherlands

www.tno.nl

T +31 88 866 00 00

TNO report

TNO 2018 R10394 | 1.0

5G and Net Neutrality: a functional analysis to feed the policy discussion

Date	13 April 2018
Author(s)	Dr P.A. Nooren, Dr N.W. Keesmaat, A.H. van den Ende, A.H.J. Norp
Copy no	
No. of copies	
Number of pages	86 (incl. appendices)
Number of appendices	3
Sponsor	Ministry of Economic and Climate Affairs, KPN, T-Mobile, FME (on behalf of Ericsson, Nokia, Huawei)
Project name	5G and Net Neutrality
Project number	060.23412

The analysis set out in this report is that of the authors and does not necessarily reflect the official opinion of the project co-sponsors.

Management Summary

Situation

The importance of mobile connectivity grows as networks and applications expand further in important sectors in society, such as mobility and transport, health, manufacturing, media and public safety. Many of the applications in these so-called verticals are expected to demand tailored mobile connectivity, for example extremely short delays, high reliability or low power consumption. With today's 4G mobile networks, it proves to be difficult for mobile network operators to meet such demands, for example, because of technical limitations or high costs. The next generation of mobile networks, commonly labelled as 5G, is developed to address the new application requirements through technologies such as network slicing and edge computing. With these new technologies, mobile operators have the technical capability to provide a range of connectivity flavours, tailored to the diverse requirements from vertical applications.

Mobile networks, including those with the new 5G connectivity features, are subject to the rules in the EU regulatory framework, The EU Regulation 2015/2120 sets the rules for net neutrality. BEREC has published Guidelines that provide guidance on the implementation of the rules. The Regulation and the Guidelines emphasize the open access of consumers to the global public internet. To this end, they contain detailed rules and guidance aimed at protecting Internet Access Services. The general rule is that Internet Service Providers (ISP)s must treat all traffic equally, which seems to be at odds with the 5G view to provide tailored connectivity to verticals and applications. In a further refinement of the general rule, the Regulation and Guidelines do offer room for traffic management and differentiation between traffic flows, subject to specific conditions. There is also the option to provide so-called Specialised Services in parallel to Internet Access Services, again subject to specific conditions.

The views among policymakers and industry on the alignment of the EU Regulation with its rules and conditions for Internet Access Services, Specialised Services, and 5G technology vary and have led to debate:

- Several industry parties fear a strict interpretation of the rules, which would in their view prevent the roll-out of tailored network services for verticals and reduce 5G to merely a faster version of 4G;
- Several policymakers expect that the Regulation and Guidelines provide the room needed for the uptake of a range of differentiated IP connectivity services and therefore cannot readily acknowledge these industry concerns.

The different views introduce a degree of uncertainty on what types of tailored connectivity will be allowed in 5G networks. This uncertainty can affect the technical and investment roadmaps of the operators and the companies in sector verticals. Industry parties and policymakers do, however, agree on the overall need for the roll-out of 5G infrastructure and applications, for business and societal reasons.

Approach and scope

TNO has taken the initiative for this study that aims at providing a functional and factual analysis of the alignment between 5G and net neutrality. The study is motivated by earlier discussions that TNO had with the Ministry of Economic Affairs and Climate Policy, the Authority for Consumers and Markets, telecom operators KPN and T-Mobile and equipment suppliers Nokia, Ericsson and Huawei (through the industry association FME). The study approach has been proposed by TNO and accepted by these sponsors. The assessment of the alignment of 5G with European net neutrality rules is carried out according to the following steps:

- Identification and description of key connectivity requirements of future applications in three sectors selected by TNO: Media, Intelligent Transport Systems and Public Safety;
- Identification and description of the key technical options in future mobile networks for providing such connectivity, based on the 5G network functions that are being standardized by 3GPP;
- Mapping of the European Net Neutrality Regulation and Guidelines to these options, in the context of the selected application domains;
- Assessment of the alignment between the 5G architecture options and the net neutrality rules, including the indication of areas where the application of the rules is expected to be relatively straightforward and where their application can be expected to be more complex.

The analysis has been restricted to the technical description of mobile connectivity required in emerging applications and the mapping of net neutrality rules to this connectivity. This means that business and commercial aspects, the formulation of policy recommendations and suggestions for changes to the Regulation and Guidelines are explicitly out of scope of the project. The analysis has been conducted using publicly available and verifiable sources. The technical analysis of 5G technology is based on 3GPP Release 15 specifications. The net neutrality rules and their interpretation are taken from the EU Regulation and the BEREC Guidelines. In addition to these sources, we have benefitted from the information and insights provided by subject matter experts in a series of interviews.

Three use cases

For the identification and description of the key connectivity requirements in the sector verticals, three specific use cases have been developed, one for each sector. The use cases obviously have a narrower scope than the sectors they are taken from. Still, each of them introduces crucial connectivity requirements. Together, they present a variety of challenging requirements for 5G mobile networks. The use cases are:

- *Virtual Reality (VR) in media and entertainment*. The next generation of VR applications builds on the availability and growing adoption of head-mounted devices like the Samsung Gear VR and Oculus Rift. The streaming of 360-degree VR content introduces challenging requirements for bandwidth and network latency. The VR case is also relevant because of the potentially large impact on the overall network load in case of mass market adoption.

- *Critical communications in Public Safety.* Reliable mobile communications are crucial for the effective operation of police, fire brigade and medical services during emergency situations. Until now, dedicated networks based on the TETRA standard have been used to guarantee the high service availability requirements in a broad variety of calamity scenarios. The public safety sector has recognised the need to move from dedicated standards to generic commercial technology for their critical communications. This introduces a very stringent requirement for the availability and reliability of mobile connectivity.
- *Automated Driving.* In automated driving, vehicles will maintain a certain required level of autonomy but also make use of sophisticated cloud services. Enhanced driving and manoeuvring functions typically require an environmental perception beyond the vehicle's own sensor range, such as positions and speeds of other vehicles and traffic light systems. Automated driving applications introduce stringent requirements for the reliability of the connectivity. Depending on the specific automotive function, the required network latency must be very low.

5G technology ingredients

3GPP has set several goals for its development of 5G, such as the support of higher data rates, larger network capacities and a (much) higher number of devices. Another an important goal is to introduce the technical capability for mobile operators to provide tailored connectivity to specific sectors, user groups and applications. This goal is crucial in the context of this study and it is reflected in the following key 5G technology ingredients:

- *Network Slicing*. Through (network) slicing, mobile operators can create separated virtual mobile networks on top of a single physical network infrastructure, both in the radio and the core network. Different slices can have different performance characteristics, for example in bandwidth, latency, reliability and the types and numbers of devices they can handle. Slices can also contain specific processing and storage functions.
- *Local access to Data Networks and Edge Computing*. Local access architectures aim to improve the latency and bandwidths offered to end users and applications by shortening the distance that traffic travels in the mobile network. This is done by handing over the traffic to the internet or to application servers near the location of the end user.
- *QoS differentiation*. QoS differentiation in 5G is to a large extent similar to that in 4G. It enables mobile operators to differentiate between traffic flows and introduce relative priorities. In 3GPP, several 5G QoS Identifier values have been standardised with an indication of example services for which they could be used, such as voice, real-time gaming and mission-critical data.
- *Unified access control*. Access control provides a mechanism for mobile operators to bar or allow network access for selected categories of devices. It is aimed providing coarse-grained traffic management during severe congestion situations. Examples are the option to provide access for emergency calls only or only for devices configured for mission critical services.

For each of the three use cases in this study, various combinations of the 5G technology ingredients have been used to develop several options for their implementation in 5G architectures. In a second step, the architecture options for the three use cases have been consolidated in a single 5G architecture model. This consolidated 5G architecture is the technical starting point for the assessment of the alignment of 5G with net neutrality rules.

Conclusions on alignment of 5G architectures with net neutrality rules

The technological neutrality of the Regulation allows 5G network technology itself to develop. There is no a priori ban on any 5G technology ingredient.

Our analysis underlines the importance of technological neutrality. This is a well-established principle that is adhered to in the Regulation and the Guidelines. It plays a crucial role in the analysis. What matters for the compliance with net neutrality rules is how the 5G technologies are used to support services and applications, rather than the technologies themselves. Therefore, the European net neutrality rules do not introduce a ban on any 5G technology ingredient, also not on the technologies that are being developed with the aim to differentiate between traffic flows and applications.

The assessment of the alignment of 5G with net neutrality rules depends not only on the 5G technologies, but also on the specific combination of services, applications and network architecture. It is not possible to come to an overall assessment with a single outcome on the alignment of 5G technology with net neutrality rules.

The central question in the assessment of the compliance with net neutrality rules is whether the services and applications supported by the 5G technology components adhere to the conditions and rules for Internet Access Services and Specialised Services, whichever are applicable. It is these conditions and rules that determine the room for mobile operators and content and application providers (including those from vertical sectors) in their use of 5G technology. In our analysis, slicing provides a relevant illustration of this point. Slicing is a key 5G technology that mobile operators may want to use in support of many different services and applications. The use of slicing will vary, as illustrated in the consolidated 5G architecture in the figure below.

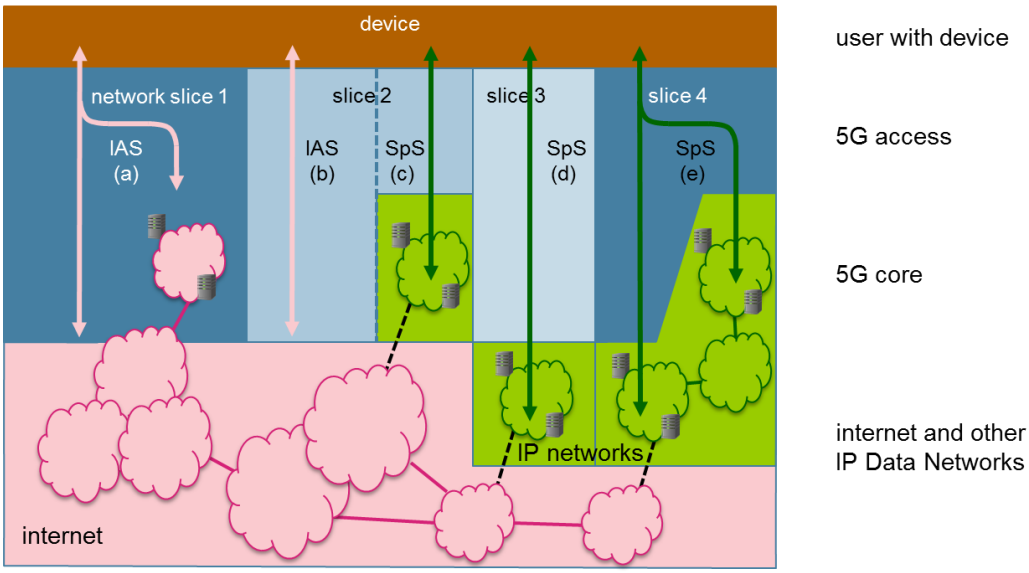

Figure: Consolidated 5G architecture view with multiple slices in a single mobile operator network, supporting Internet Access Services (IASs) and Specialised Services (SpSs).

In 5G architectures that use slicing, an Internet Access Service is always in a slice. A slice can be used exclusively to provide an Internet Access Service (slice 1). Alternatively, a single slice can be used to simultaneously provide an Internet Access Service and a Specialised Service (slice 2). A slice can also be exclusively used to provide a Specialised Service (slices 3 and 4). Thus, the use of slicing technology in a mobile operator network can bring in the rules for Internet Access Service, for Specialised Services or both, depending on the services and applications that are supported. It is not possible to come to an overall assessment with a single outcome on the alignment of slicing with net neutrality rules. This is because the topics that are encountered in the assessment and the outcome depend not only on the 5G technology, but also on the specific combination of services, applications and network architecture. This is true for network slicing, but also for other key 5G technologies such as QoS differentiation. A consequence is that mobile operators, content and application providers and national regulatory authorities will need to do further analysis to evaluate whether a particular type of (tailored) connectivity complies with the net neutrality rules.

The topics encountered in the assessment of the compliance are of varying complexity. The impact of Specialised Services on Internet Access Services *and* the objective need for optimisation in Specialised Services *are expected to have the highest complexity.*

Based on our analysis of the three use cases and the key 5G technology ingredients, we have identified nine topics that are relevant in the assessment. We have positioned these topics in the consolidated 5G architecture to show typical situations where they come into play, see the figure below.

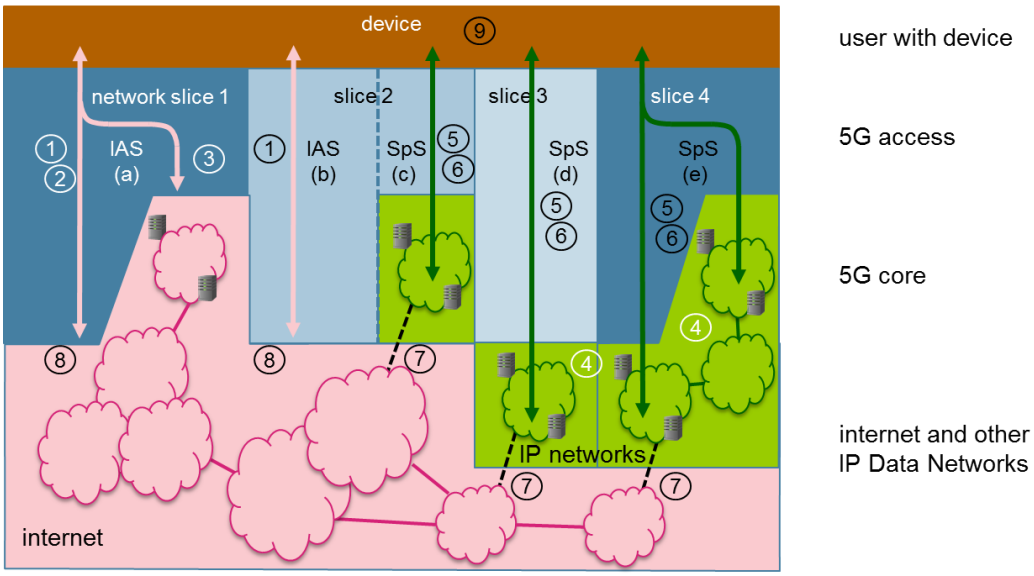

Figure: Consolidated 5G architecture with multiple slices in a single mobile operator network, supporting Internet Access Services (IASs) and Specialised Services (SpSs). The numbers indicate topics where the alignment between net neutrality rules and 5G architecture options has been investigated.

The topics are summarised in the table below, together with our expectation for their relative complexity in assessments of compliance with net neutrality rules. We define this as the relative complexity expected to be encountered by national regulatory authorities, mobile operators, and content and application providers when they analyse specific cases with more context information and (quantitative) details than the use case-inspired analysis made here. All key points mentioned in the second column in the table are discussed in the main body of this report.

Table: Topics encountered in assessment of alignment of 5G architecture options with net neutrality rules and their expected relative complexity. IAS means Internet Access Service, SpS means Specialised Service.

Topic	**Key points identified in analysis**	**Relative regulatory complexity**
1. Multiple IASs with different traffic management settings (section **5.4**)	• Interpretation of *sender and receiver* in Art 3.3 of the Regulation • Note: assumption needed in remainder of analysis - it is allowed to have multiple IASs with different traffic management settings for a given end user	low
2. QoS differentiation within IAS (section **5.5**)	• Applications with multiple different traffic flows • Transparency through standardised traffic classes or other methods • Dependency of ISP on other entities for assignment of traffic flows to traffic categories • Duration of QoS differentiation	medium to high
3. Local access to the internet (section **5.6**)	• (potentially:) IP interconnection of local networks	low
4. Public and private services and associated networks (section **5.7**)	• Size and scope of predetermined group of end users in private service	low to medium
5. Objective need for optimisation in SpS (section **5.8**)	• Determination of IAS for benchmark in case of multiple IAS offers • Variation of IAS performance between geographical regions and operators • Services comprising multiple traffic flows	high, except if SpS requirements are clearly much stricter than achievable over IAS.
6. Impact of SpS on IASs (section **5.9**)	• Multiple IASs affected by one SpS, within and outside the slice used for the SpS. • Isolation of the effect of the SpS on IAS from other effects occurring in mobile network at the same time • Complexity of network and capacity management in mobile network with many services and applications in general	high
7. SpS and connections to the internet (section **5.10**)	• Connectivity to internet from SpS through separate IAS • Connectivity between different legs between end user device and internet	low
8. Connectivity to limited number of internet end points (section **5.11**)	• Evaluation whether sub-internet service is acceptable for providing connectivity in specific situations	medium
9. Access control (section **5.12**)	(no issues if use is restricted to network congestion in emergency situations)	low

In our analysis, we found that several topics which appear to be complex at first sight, such as *Specialised Service and connections to the internet*, become relatively straightforward to assess once the details of the architecture options and the Regulation and the Guidelines are carefully combined. We expect that the low to medium complexity topics lend themselves to the formulation of “rules of thumb” within national regulatory authorities, mobile operators, and content and application providers. They can be formulated based on internal analysis or, at a later stage, be derived from the outcomes of earlier cases assessed by national regulatory authorities. Other topics such as the *impact of Specialised Service on Internet Access Services* can be expected to remain relatively complex. There are no fundamental problems that prohibit their analysis. However, the complexity of these topics is likely to make them unsuitable for a generic “rule of thumb” approaches. They require a case-by-case approach. The complexity depends on the level of detail that national regulatory authorities, mobile operators, and content and application providers pursue in their analyses.

The topics encountered in the assessment are relevant for services and applications provided over mobile and fixed networks in general. They are not exclusively related to 5G technology.

A final observation is that the topics identified as relevant in the assessment are not exclusively related to 5G. They can also present themselves in the analysis of services and applications provided over 3G, 4G and pre-5G networks. As the Regulation and Guidelines are to a (very) large extent technology neutral, the analysis of the topics would be largely similar. The topics can be expected to be more relevant in 5G networks though, as 5G technology provides more extensive support and flexibility for tailored mobile connectivity aimed at specific sectors or user groups. The topics can also present them in fixed networks.

Recommendations

Our first recommendation is to clearly distinguish between 5G architecture elements on the one hand and the net neutrality concepts of Internet Access Service and Specialised Service on the other. One should keep a technology-neutral view and not attempt to define a one-to-one mapping between the two. Two important examples of this are:

1. A slice is not the same as a specialised service. Slicing can be used to support an Internet Access Service, a Specialised Service or both.
2. The application of QoS differentiation is not limited to Internet Access Service. QoS differentiation can be used as a method for traffic management within an Internet Access Service. However, it can also be used to assure the quality of Specialised Services. A prominent example of the latter is the VoLTE architecture in 4G networks.

Our second recommendation is that subject matter experts at national regulatory authorities, mobile operators, and content and application providers build upon our approach and findings in their assessments. We expect that the consolidated architecture model provides a good starting point to structure the overall discussion on services and applications over 5G networks and their compliance with net

neutrality rules. For the analysis of specific services and applications, the three-step approach applied to the use cases in this report is recommended:

1. Determine the connectivity requirements of the services and applications in the use case.
2. Develop the 5G architecture options to support the connectivity requirements. The 5G technology ingredients described in this report are expected to play an important role here.
3. Evaluate the alignment of the combination of services, applications and architecture options with net neutrality rules. Here, the analysis of the specific topics made in this report can probably (partly) be reused.

Mobile operators, content and application providers and national regulatory authorities can use this approach to develop their own individual analysis. These steps can also be used to structure the discussion among stakeholders and come to a shared analysis. Such a shared analysis would be beneficial for providing clarity and reducing uncertainties that industry may encounter in its development of roadmaps for 5G networks and applications that rely on tailored connectivity.

Contents

Appendices

1 Introduction

1.1 Situation

The importance of mobile connectivity grows as networks and applications expand further in important sectors in society, such as mobility and transport, health, manufacturing, media and public safety. Many of the applications in these so-called verticals are expected to demand tailored mobile connectivity, for example extremely short delays, high reliability or low power consumption. With today's 4G mobile networks, it proves to be difficult for mobile network operators to meet such demands, for example, because of technical limitations or high costs. The next generation of mobile networks, commonly labelled as "5G", holds the promise that capabilities and features will be inherently built in which enable Mobile Network Operators (MNOs) to offer truly differentiated services to different sectors and market segments. This is also expected to open up new revenue streams. In this sense, 5G is not meant to be just a straightforward follow-up of 4G, but it is expected to bring something new to the scene. As illustrated in Figure 1, the design of 5G networks is targeted towards three main service categories: enhanced Mobile Broadband (eMBB), ultra-reliable and low-latency communications (uRLLC) and massive Machine Type Communications (mMTC). Together, these are expected to stimulate the development and adoption of various sector-specific mobile services and applications. With new technologies and functions such as Edge Computing and Slicing, mobile networks can provide a range of connectivity flavours.

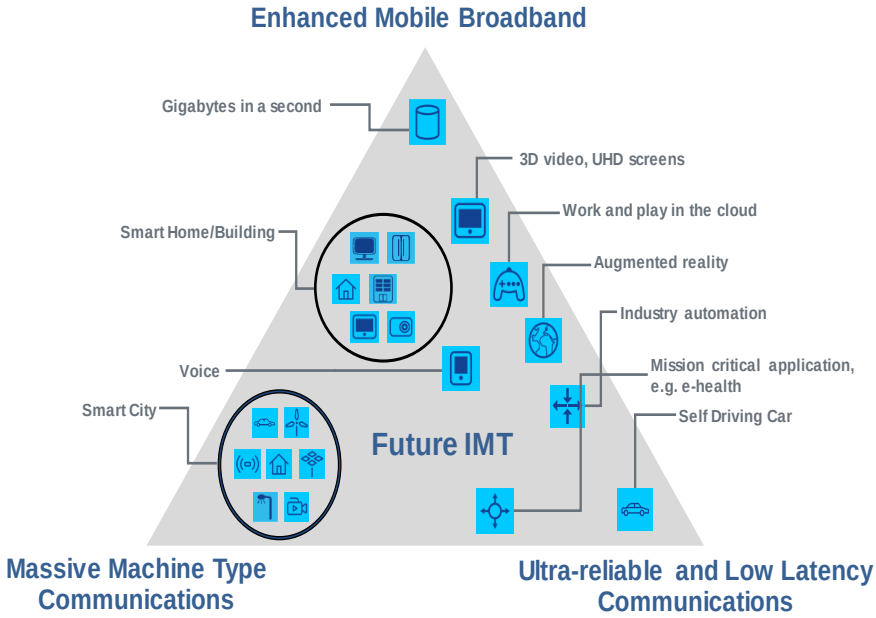

Figure 1: Three main service categories targeted in 5G (Source: ITU-R M.2083-0 [1]).

Mobile networks, including those with the new 5G connectivity features, are subject to the rules in the EU regulatory framework for electronic communications, including specific rules on net neutrality. The EU Regulation 2015/2120 [2] sets the rules for net neutrality. BEREC has published Guidelines [3] that provide guidance on the regulatory implementation of the rules. The Regulation and the Guidelines

emphasize the open access of consumers to the global public internet. The views among policymakers and industry on the alignment of the EU Regulation and (still to be deployed) 5G technology vary and have led to the following debate:

- Several industry parties fear a strict interpretation of the rules, which would in their view prevent the roll-out of tailored network services for verticals and reduce 5G to merely a faster version of 4G;
- Several policymakers expect that the Regulation and Guidelines provide the room needed for the uptake of a range of differentiated IP connectivity services and therefore cannot readily acknowledge these concerns.

The different views introduce a degree of uncertainty on what types of tailored connectivity will be allowed in 5G networks. This uncertainty can affect the technical and investment roadmaps of the operators and the companies in sector verticals. Industry parties and policymakers do, however, agree on the overall need for the roll-out of 5G infrastructure and applications, for business and societal reasons.

1.2 Motivation for the 5G and Net Neutrality project

In 2016 and 2017, TNO noticed that the discussions on the alignment of 5G and net neutrality tended to be restricted to the exchange of opinions at a relatively high level. Subsequent talks with the Ministry of Economic Affairs and Climate Policy, the Authority for Consumers and Markets, telecom operators KPN and T-Mobile and equipment suppliers Nokia, Ericsson and Huawei (through the industry association FME) confirmed our assumption that the discussion would benefit from a solid analytical and factual underpinning. Such an underpinning would be useful in at least two areas: the mobile connectivity that will be needed by the verticals in specific cases, and the technical options in future mobile networks to provide such connectivity.

This motivated TNO to formulate a project to develop a functional and factual analysis of 5G and net neutrality. Apart from the analysis of the requirements from use cases and mobile connectivity options, the project includes the mapping between the Regulation and BEREC Guidelines and the applications and mobile connectivity. The mapping should identify (1) the areas where the 5G technology options and the net neutrality rules are in alignment, (2) areas where they are no aligned and (3) areas where the mapping cannot be completed because of multiple potential interpretations.

1.3 Scope of the analysis

The scope of the project was proposed by TNO to the aforementioned stakeholders and forms the basis of this study:

- identification and description of key connectivity requirements of future applications in three sectors selected by TNO: Media, Intelligent Transport Systems (ITS) and Public Safety;
- identification and description of the key technical options in future mobile networks for providing such connectivity, based on the 5G network functions that are being standardized by 3GPP;

- mapping of the European Net Neutrality Regulation and Guidelines to these options, in the context of the selected application domains;
- assessment of the alignment between the 5G architecture options and the net neutrality rules, including the indication of areas where the application of the rules is expected to be relatively straightforward and those where their application can be expected to be more complex.

The analysis has been restricted to the technical description of mobile connectivity required in emerging applications and the mapping of net neutrality rules to this connectivity. This means that the following aspects are outside of the scope of this study:

- business and commercial aspects[1];
- formulation of policy recommendations and suggestions for changes to the Regulation and Guidelines.

In TNO's view, industry and policymakers are in the best position to further develop these aspects. This project aims to provide stakeholders with factual and unbiased underpinning they can use in their considerations, in a way that they see fit. On the networking technology side, the scope is restricted to mobile access over 5G radio access networks. The integration of fixed access networks and satellite networks in 5G is not considered.

1.4 Approach and sources

The project approach is depicted in the figure below.

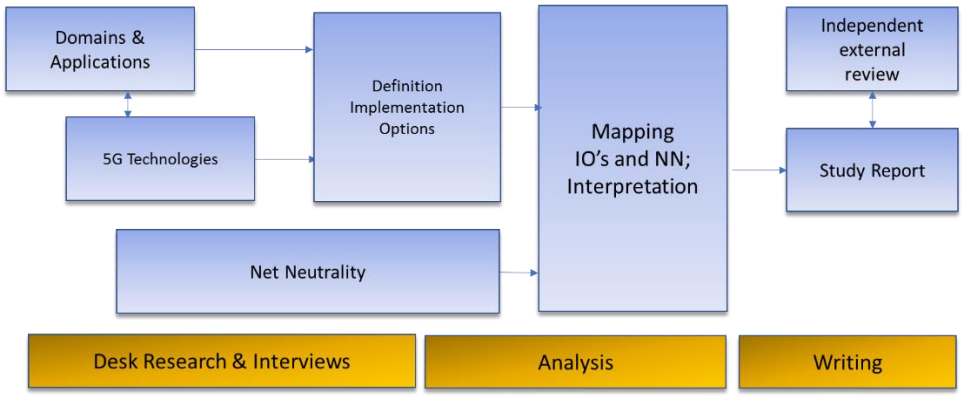

Figure 2: Approach followed in this study project.

At the start of the project, TNO selected the three vertical application domains to be analysed, based on the expected relevance of (tailored) mobile connectivity and the availability of domain knowledge at TNO. For each vertical, a specific application was selected and analysed in more detail, based on its expected future sector relevance and its dependence on 5G for certain connectivity requirements. In parallel, 5G technology ingredients were selected and described, with an emphasis on functionalities which were deemed particularly relevant for the applications

[1] An example of a commercial aspect that is addressed in the Regulation and Guidelines is the so-called zero-rating, which led to an extensive debate in the Netherlands.

chosen. Then, for each application, several 5G architecture options have been derived. These options are examples of how the application could be realized using 5G network functionalities and do not limit the actual implementation space, nor impose any specific implementation approach. In the next step, the focal point of this project is reached: the set of 5G architecture options is combined with the net neutrality rules from the Regulation and their (further) interpretation in the Guidelines. The findings have been documented in this report and reviewed by two external experts.

The analysis has been conducted using publicly available and verifiable sources:

- The technical analysis of 5G technology is based on 3GPP Release 15 specifications.
- The net neutrality rules and their interpretation are taken from the EU Regulation and the BEREC Guidelines.

In addition to these sources, we have benefitted from the information and insights provided by subject matter experts in a series of (telephone) interviews, see annex A.

1.5 Target audience

This report is targeted at subject matter experts in industry (e.g., at mobile operators, vendors and content and application providers in vertical sectors) and in government (e.g., policymakers at the EU and national level, and at NRAs).

1.6 Guide to this report

Chapter 2 of this report provides the reader with the context of the study. It sets the scene in terms of 5G mobile communications and net neutrality. Chapter 3 introduces the selected vertical domains and the use cases centred around challenging applications. The 5G technology options are described in Chapter 4. This chapter also provides a consolidated 5G architecture model that is used to position 5G implementation options. Chapter 5 discusses the alignment of 5G services, applications and architectures with the rules for net neutrality. Chapter 6 presents our conclusions and recommendations.

2 Setting the scene

This chapter presents the context for the analysis of 5G and net neutrality. Starting from an overview of mobile network evolution, 5G and its key technology ingredients are introduced. Then, after a short review of the historical background of net neutrality, the regulatory starting points for the analysis are described by highlighting the key aspects in the Regulation and the Guidelines.

2.1 Evolution of mobile networks

Mobile networks have developed in generations from 1G to 2G to 3G to 4G and will in the future reach its 5th Generation (5G). In this section the various generations will be briefly discussed, the reader is referred to other texts (e.g., [4],[5]) for more details.

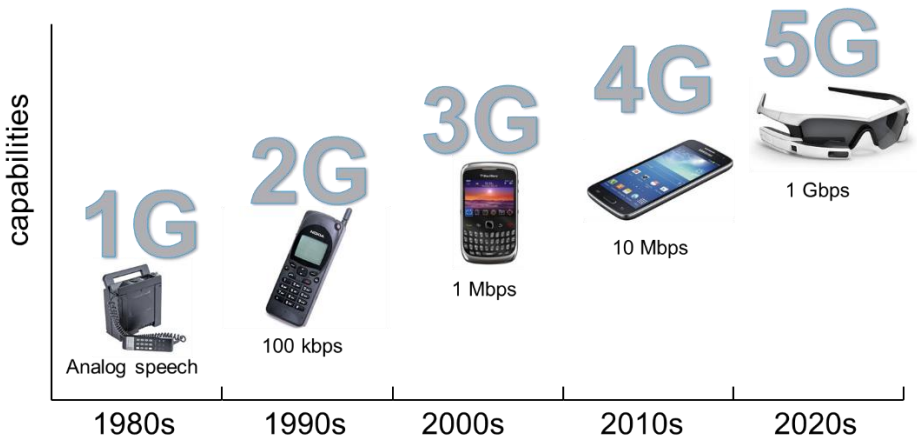

Figure 3: Development of mobile communications since late seventies. Source: TNO

1G-3G

The early generations of mobile networks used analogue technology. In the Netherlands, ATF-1[2] was introduced in 1980, ATF-2 in 1985, and ATF-3 in 1989. ATF-2 was the first cellular network based on the NMT (Nordic Mobile Telephone) standard operating at 450 MHz. ATF-3 was based on NMT-900 see [6]. NMT was also used in Belgium and Luxembourg, while other European countries deployed national standards. AMPS was deployed in the USA, as well as some countries in South-America and Asia. ETACS and NTACS were deployed in Japan.
With the introduction of 2G-GSM in 1991, the transition was made to digital technology. GSM started as a European standard specified by ETSI and was adopted worldwide later. GSM was enhanced with GPRS and EDGE to support data communications. The IS-95 standard in the USA introduced a novel technology CDMA, which was later adopted in the third generation. The third generation mobile networks started on the basis of the visions, objectives, and specifications formulated by the ITU in its IMT-2000 recommendations (see e.g. [7], [8], [9]). Based on these recommendations three network technologies have been

[2] ATF is an acronym for 'autotelefoon', Dutch for car phone, as mobile phones were called at that time.

accepted as 3G, being compatible with the installed base of three main 2G standards. For Europe, this was UMTS (Universal Mobile Telecommunication Service) specified by the 3GPP. UMTS supported packet-switched data next to circuit-switched voice. It evolved to include HSDPA and HSUPA to provide higher downlink and uplink speeds. 2G and 3G use largely the same core network architecture but they differ considerably in the radio network architecture.

4G
The framework, objectives, and specifications for 4G were set by the ITU in its IMT-Advanced documents (see e.g., [10], [11]) and this led, in 3GPP, to the development of the so-called LTE (Long Term Evolution) radio network with its EPC (Evolved Packet Core) core network, together called the EPS (Evolved Packet System). This 4G network was based on packet switching technology. With the assistance of IMS (IP-Multimedia Subsystem) voice communication can be provided using Voice over LTE (VoLTE) technology. Data speeds and latency have improved a lot, and both the radio network and the core network architecture have changed considerably. The LTE radio network technology has been extended to LTE-Advanced and lately also to LTE-Advanced Pro, focussing on even higher speeds and lower latency.

5G
The development of the latest generation of mobile networks has been stimulated by the ITU IMT-2020 programme [1] and special interest groups such as NGMN (Next Generation Mobile Networks) and 5G-PPP. Research into 5G has been ongoing since 2011. 3GPP has been working on the specification of 5G since 2015. In addition to higher speeds and more capacity to keep up with the growth of mobile data, the development of this generation has also focussed on the support for various so-called verticals such as automotive, energy, health, entertainment, smart cities and smart industry. This brings in requirements for higher reliability, lower latency and a higher number of devices. The support of multiple verticals also introduces the need for diversification which has led to the development of the concept of network slicing. The three major areas identified for 5G by ITU [1] comprise: Enhanced mobile broadband (eMBB), Ultra reliable and Low latency communication (URLLC), and Massive machine type communications (mIoT), see also Figure 1. The standardisation of 5G is divided into two main phases that are related to specific 3GPP releases.

Phase 1:
This phase will result in documents for 3GPP Release 15 and has a target (freeze) date of June 2018. At the time of writing of this report, the requirements and architecture level documents have been finalised. The documents specifying the management of network slicing are still under development, as are the documents on protocols and the radio network.

Phase 1 will contain the basic features with forward compatibility, such as roaming, charging, management, QoS and policy control, service continuity, and network sharing. The focus is on enhanced Mobile Broadband and selected parts of the Ultra-Reliable and Low Latency Services aspects. This network focussed phase offers basic support for network slicing.

Phase 2:
This phase has an intended target (freeze) date of December 2019. It will focus on massive IoT and the enhancement of the aspects started in Phase 1. Other topics that may be addressed include multicast, device-to-device, vehicle-to-x, satellite support, trusted non-3GPP access, self-organizing networks, virtual LAN support, support for factory networks and support for operational railway communications.

2.2 Key 5G technology ingredients

A high-level overview of a 5G network is presented in Figure 4.

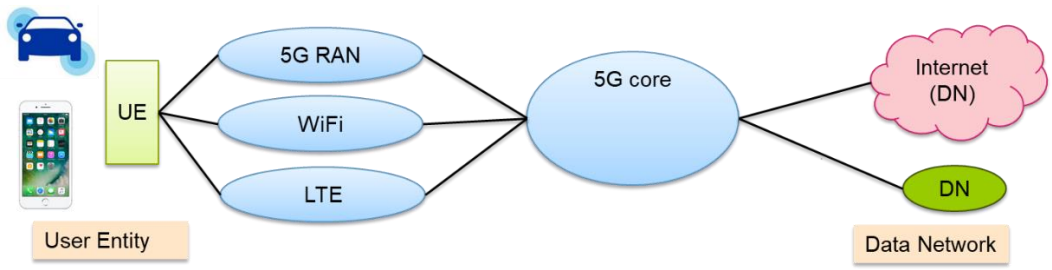

Figure 4: High-level overview of a 5G network.

The following components can be identified:[3]

- UE (User Equipment), i.e. the mobile device;
- Radio Access Networks (RAN), such as:
  - 5G RAN, i.e. the 'native' 3GPP radio access network for 5G – note that also the term NR (New Radio) is used here;
  - WiFi;
  - LTE, i.e. 4G radio networks connected to a 5G core network;
- 5G Core Network (5GC);
- DN (Data Network), a network outside of the 5G network. The internet is a prominent example of a DN but there are also other examples.

Within 3GPP, the architecture of the core network has been standardized in TS 23.501 [12]. The non-roaming architecture defined in that document is as depicted in the following diagram. It should be noted that this is a functional architecture, and in actual implementations, functions may occur more than once.

[3] For an overview and explanation of the abbreviations used see Annex B.

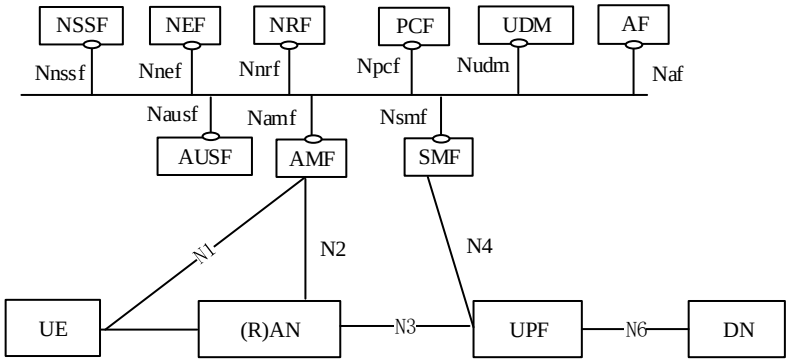

Figure 5: Non-roaming 5G architecture from TS 23.501.

In this functional network architecture, the UE and (R)AN represent the mobile device and the radio access network. The UPF is the User Plane Function carrying the user data and the DN is the Data Network that is external to the 3GPP network. The other functions represent the Control Plane and they operate via the so-called service-based interfaces. These are the AMF, the Access and Mobility Management Function, which is the first point of contact for the UE when it tries to access the mobile network. The SMF, the Session Management Function, is involved in the handling of bearers, called PDU Sessions in 5G, and controls the UPF. The AUSF, the Authentication Server Function, controls the authentication of the UEs and corresponds to the AUC in 4G networks. The UDM, the Unified Data Management, contains the subscriber data and roughly corresponds to the HSS in 4G networks. The PCF, the Policy Control Function, and the AF, the Application Function, correspond with the PCRF and AF in 4G networks and they are involved in the handling of traffic management policies and QoS. The NSSF, the Networks Slice Selection Function, is a new function in 5G involved in the handling of slices (see section 2.2.1). Finally, the NRF (Network function Repository Function), and the NEF (Network Exposure Function) enable service discovery and service exposure to 3rd parties.

As a complement to the service-based representation, TS 23.501 also provides a traditional reference-point representation of the 5G architecture. Figure 6 shows the non-roaming architecture in this representation.

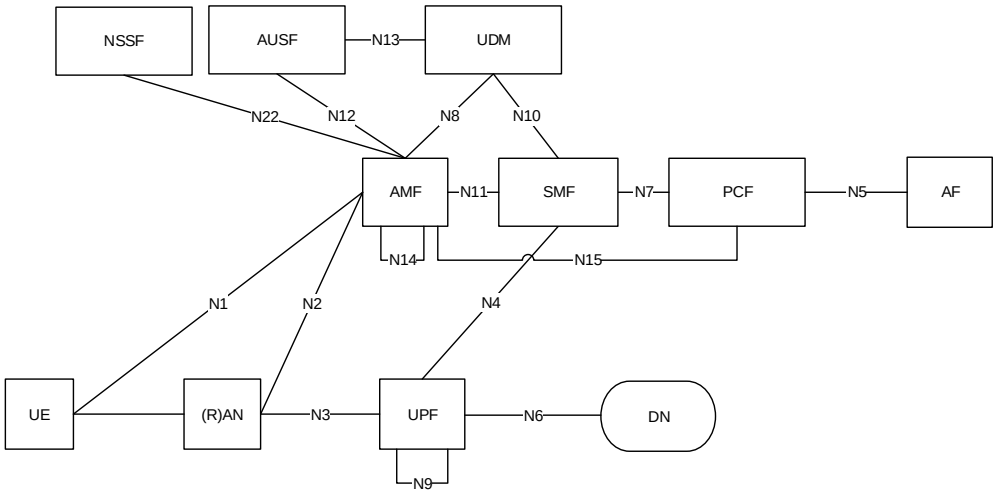

Figure 6: Non-roaming 5G architecture from TS 23.501 (reference-point representation).

In addition to the above non-roaming architecture also several roaming architectures (both the local breakout version and the home routed version) have been defined, as well as several architectures for local access (see section 2.2.2).

In the next sections, four key components of the 5G architecture will be elaborated:

1. (Network) Slicing;
2. Local Access with or without Service Hosting;
3. Quality of Service (QoS), and QoS Differentiation;
4. Access Control and Access Barring.

These components are directly related to the technical capability for mobile operators to provide tailored connectivity to specific sectors user groups and applications. This makes them very relevant for the discussion of the three use cases included in this study,

2.2.1 Network slicing

Network slicing has been specified in 3GPP in various normative documents both on the requirement level (TS 22.261 [13]), on the architecture level (TS 23.501) and at the management level (TS 28.530 [14]). Network slicing is seen as a method to enable operators to handle a great variety of requirements for different verticals. It also enables logical separation of different types of networks (e.g. an eMBB network separate from a URLLC network, see section 2.1). It is foreseen that network slices are based on virtualized networking enabled by well-known virtualization techniques.

A high-level diagram depicting the concept of slices is shown in Figure 7.

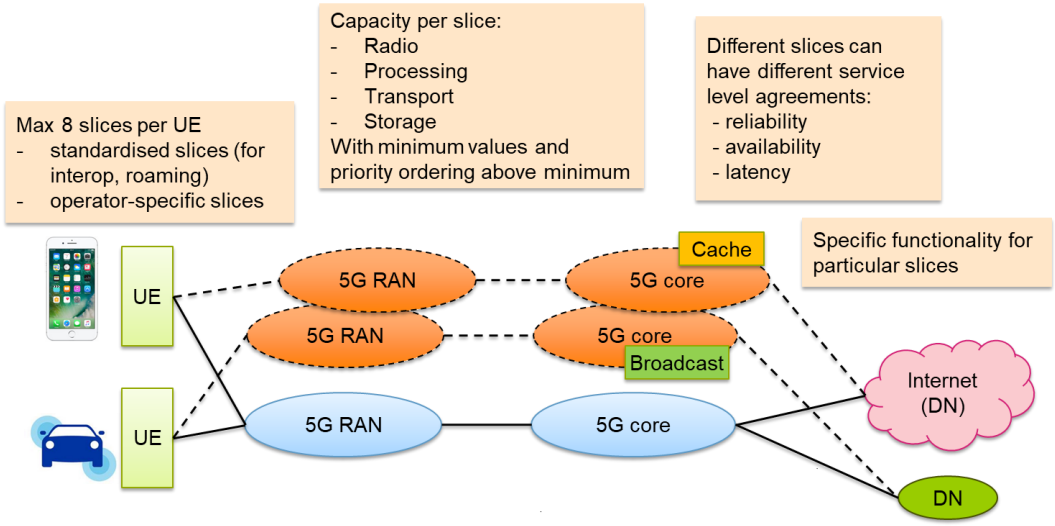

Figure 7: Network slicing in 5G RAN and core networks.

Network slicing is defined for 5G core networks where each slice can contain certain network functions with varying performance capabilities and it is being defined for 5G radio networks, where radio network nodes (the so-called gNodeBs or gNBs) are made slice-aware. A device (called User Entity or UE in 3GPP) can select slices for its services, but it will not be able to use more than 8 slices simultaneously. The number of slices provided by operators is not limited. Two pictorial representations of slices are given in Figure 8 and Figure 9. Figure 8 illustrates the various Network Functions and their interfaces. Note that Network Functions can be 'inside' a specific slice or be part of multiple or even all slices. Figure 9 illustrates that the slice arrangement in the RAN can be different from the Core Network and consequently different slice combinations can be made for an end-to-end connection (slice stitching).

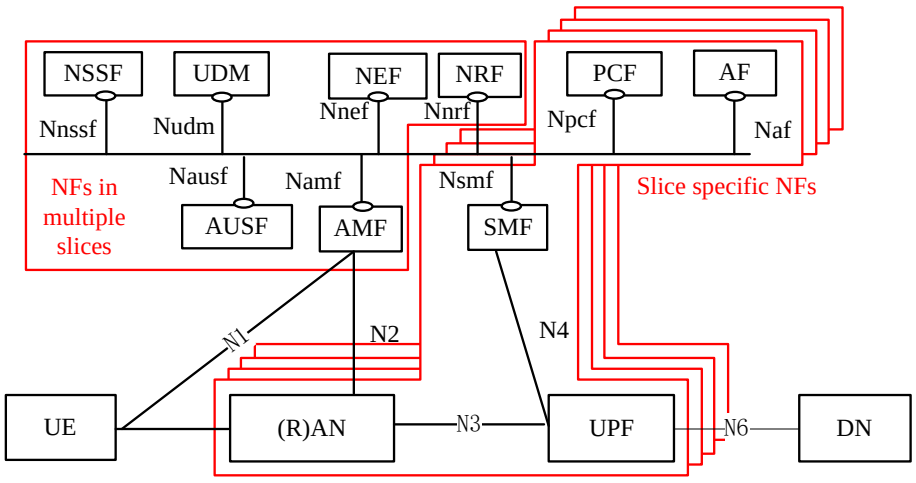

Figure 8: Network functions can be in only one or in multiple slices. Source: TNO, based on 3GPP TS.23.501 [10].

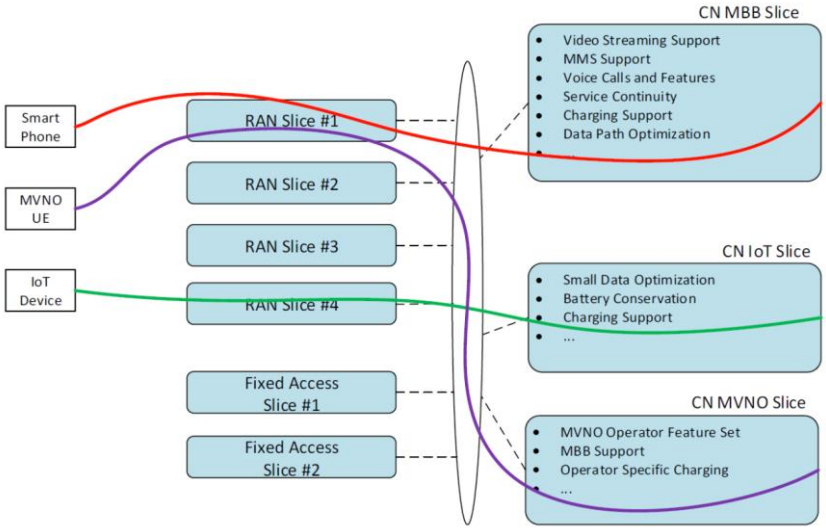

Figure 9: Options for combinations of RAN and core slices, source: 5G Americas [15].

At the requirements level (TS 22.261), slices can have a minimum and maximum capacity and a priority order between slices that becomes relevant in case of competition for resources. At architecture level (TS 23.501), these requirements particularly dealing with capacity management have not yet led to specific functionality. At management level (TS 28.530) this requirement may result in some functionality, but at the current state of specification, no specific functionality has been formulated yet. The corresponding non-normative study (TS 28.801 [16]) provides a use case on the relation between slices.

At architectural level (TS 23.501), slice functionality mainly concerns aspects such as identification and selection of a Network Slice through an information element called S-NSSAI (Single Network Slice Selection Assistance Information), but also subscription aspects, UE configuration aspects, roaming support and interworking with 4G. Slice selection is also related to the process of getting access to a network. Before slice selection can start, an initial network access is required, hence an initial AMF function is used for getting access. After subsequent slice selection, the initial AMF function may be exchanged for another AMF that is applicable to the selected slice [17].

At the time of writing of this report, three standardized Slice Service Types (SSTs) are specified, for eMBB, for URLLC and for MIoT. Apart from these, operators may define their own slice types, and use further Slice Differentiators (SDs) to distinguish between slices of the same slice service type.

2.2.2 Local access to DN and Edge Computing

Local access has been specified in 3GPP to achieve some of the high-performance requirements as part of 5G, such a throughput and latency. Terminology related to local access is not always consistently used and mixed with other terms, so it is necessary to clarify the following terms:

- **Local Breakout**: this term should not be confused with Local Access; Local Breakout is used in 3GPP only in the context of roaming, where it is used as the opposite of home routing.
- **Mobile/Multi-access Edge Computing, MEC**: this term is not used in 3GPP; it is a term defined by ETSI and is related with Local Access, but ETSI has defined it largely with 4G networks in mind. It is also identified as a relevant technology by 5G Americas [18];
- **Edge computing:** this term is used a few times at the architecture level (TS 23.501) in 3GPP, but it mainly refers to the use of Hosted Services;
- **Hosted Services:** this term is used mainly at the requirements level (TS 22.261) and refers to services containing an operator's own applications or trusted 3rd party applications. Hosted Services are offered via a Service Hosting Environment;
- **Service Hosting Environment:** this term is defined at the requirements level (TS 22.261) and denotes an environment fully controlled by the operator from which to offer Hosted Services;
- **Local access (to Data Network)/Distributed User Plane Function:** these terms are not defined but they are used at the architectural level (TS 23.501) in 3GPP; local access refers to routing of traffic to a Data Network at or near the edge/access of the mobile network; distributed UPF "located close to or at the Access Network site" is contrasted to centrally located UPF.

In this report, we use the term Local Access, and this can be used with or without a Service Hosting Environment. Local Access without Service Hosting Environment corresponds to exiting the 3GPP network locally to a non-operator-controlled Data Network (e.g. the internet). Local Access with Service Hosting Environment corresponds to exiting the 3GPP network locally to an operator-controlled network from which operator or trusted 3rd party services can be provided, see Figure 10.

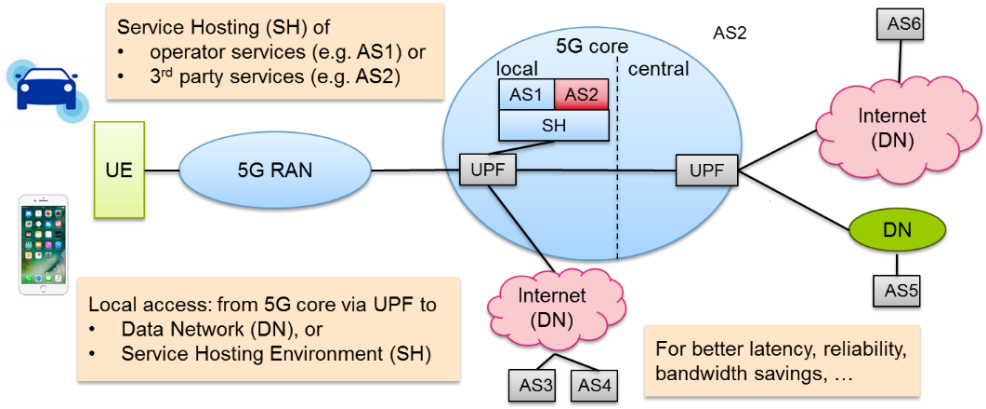

Figure 10: Local access with and without service hosting in a 5G architecture.

The specification of local access is based on the architecture depicted in Figure 11.

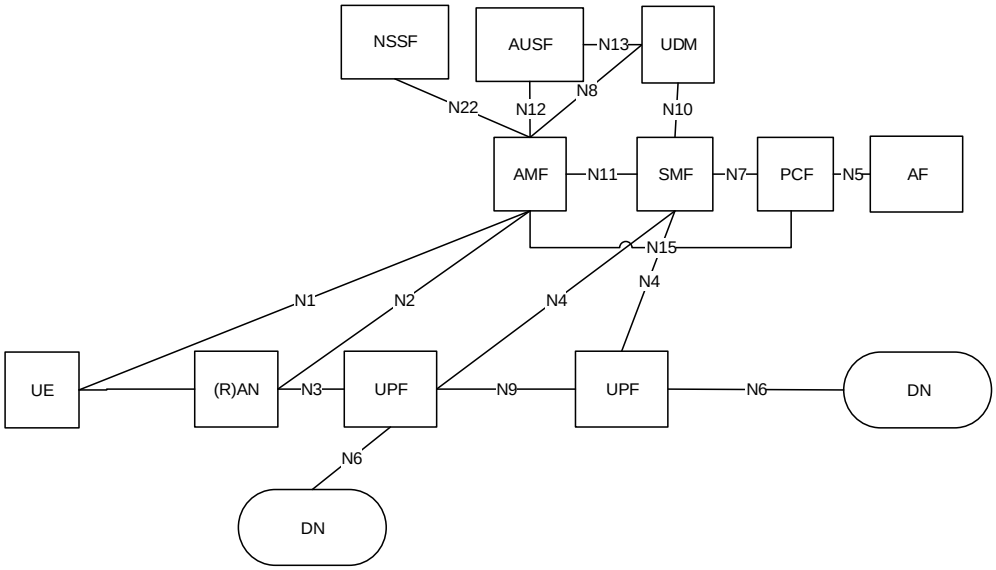

Figure 11: Specification of Local Access in 5G architecture

Significant in this diagram is the use of multiple (chained) UPFs, where the leftmost UPF is used to provide local access to the local DN, whereas the rightmost UPF provides access to the central DN to the right. The SMF function controls the UPFs and instructs them how and when to route traffic locally and when to route traffic centrally. Technically there are different solutions depending on the use of IPv4 or IPv6, where IPv6 is more flexible in handling multiple exit points.

2.2.3 QoS differentiation

In 5G networks, the options and functions for QoS differentiation are largely similar to those in 4G networks. A single UE may have multiple bearers, called QoS Flows in 5G. The QoS flows of multiple UEs share the radio capacity in the radio cell.

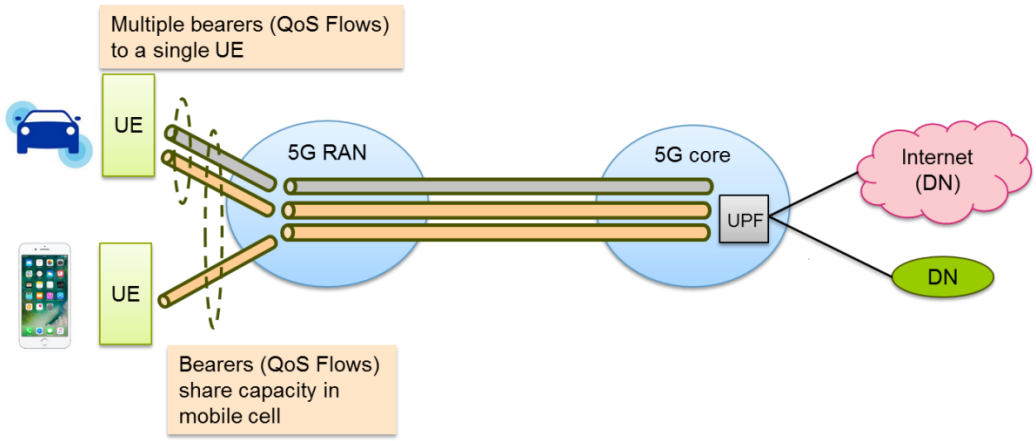

Figure 12: QoS differentiation in a 5G architecture.

In 5G networks, the terminology related to QoS has changed and some new parameters and QoS values have been defined, mainly related to the extreme low latency QoS cases. In this section, an overview of 5G QoS is given. More details are provided in Annex C.

- In 5G networks, a QoS Flow is identified by a *QoS Flow ID* (QFI) and refers to a *QoS Profile*. A QoS Profile contains *QoS parameters* comprising among others a *5G QoS Identifier* (5QI), an *Allocation and Retention Priority* (ARP), and some other parameters (e.g. related to guaranteed, maximum and aggregated bit rate).
- For the standardized 5QI values, a mapping is specified on *QoS characteristics* such as (maximum) packet delay and (maximum) packet loss. The QoS characteristics are guidelines for the network element and base station settings. In 5G networks, it is also possible to signal QoS characteristics themselves without direct correspondence to a standardized 5QI.
- As in 4G networks, the ARP parameter determines which QoS Flows can be pre-empted (i.e. can be 'pushed' from the network), and which QoS Flows are capable of pre-emption (i.e. can 'push' other flows from the network).

TS 23.501 provides a table that maps 5QI values to QoS Characteristics. An extract is provided below in Table 1, the full table is contained in Annex C.

Table 1. Mapping of 5QI values to QoS Characteristics (Extract of Table 5.7.4-1 from TS 23.501)

5QI Value	Resource Type	Packet Delay Budget	Packet Error Rate	Example Services
B	Delay Critical GBR	5 ms	10^{-5}	Remote control
1	GBR	100 ms	10^{-2}	Conversational Voice
2		150 ms	10^{-3}	Conversational Video (Live Streaming)
65		75 ms	10^{-2}	Mission Critical user plane Push To Talk voice (e.g., MCPTT)
5	Non-GBR	100 ms	10^{-6}	IMS Signalling
6		300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)

The 5QI values with extreme low latencies (B-G) are new in 5G. Note that – as in 4G networks –separate 5QI values are used for Mission Critical services (65, 66, 69, 70, 75, 79).

In current 4G networks, one prominent example of the use of QoS is provided by the VoLTE (Voice over LTE) and ViLTE (Video over LTE) services. For the IMS signalling bearer the (4G) QCI = 5 is prescribed by GSMA [19], for the voice bearer (in VoLTE) QCI = 1 is prescribed by GSMA [19] and for the video bearer (in VoLTE) QCI = 2 is prescribed by GSMA [20]. Standard internet access usually uses QCI = 6, but this is not standardized.

As explained above, the 3GPP architecture expects the underlying networks and base stations to ensure the required QoS characteristics (such as packet delay, packet loss) without specifying how. In IP networks, QoS mechanisms such as DiffServ exist for QoS Differentiation, but these are not widely used end-to-end in public networks. In base stations, the QoS characteristics can be reflected in the scheduling of data packets or in the use of different radio technology options (such as switching on/off error correction).

Similar to 4G, 5G will support both IPv4 and IPv6. 5G will also support QoS flows with Ethernet or so-called unstructured data. These can be used to connect a mobile device to, for example, a factory Ethernet network or a company LAN. QoS differentiation applies to non-IP QoS flows in the same way as to the IPv4 and IPv6 flows.

2.2.4 Unified access control

Networks access control already exists in 4G networks, through a combination of several mechanisms (such as access class barring, extended access class barring, and congestion control). For 5G networks, a new Unified Access Control has been standardised. Figure 12 shows the network functionality underlying this access control.

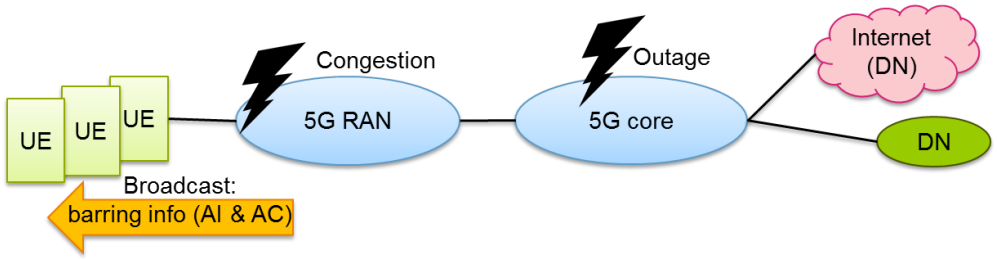

Figure 13: Unified access control in a 5G architecture.

Access control involves the broadcasting of barring information to UEs in case of congestion or outages. In the barring information, two types of information components are included: Access Identity (AI) and Access Category (AC). Based on this information, each individual UE of which the specific configuration is coded with an Access Identity number. Access identity is an indication of the type of UE to allow barring of access seeking terminals by UE type. The UE knows its own AI, so it can assess whether a broadcasted inhibit to access the network is applicable to this UE (see TS 22.261).

The list of access identities is given in Table 2. This table is included here merely for illustrative purposes and will not be explained in detail.

Table 2. Access Identity numbers used in 5G unified access control (extract from TS 22.261)

Access Identity number	UE configuration
0	UE is not configured with any parameters from this table
1	UE is configured for Multimedia Priority Service (MPS).
2	UE is configured for Mission Critical Service (MCS).
3-10	Reserved for future use
11	Access Class 11 is configured in the UE.
12	Access Class 12 is configured in the UE.
13	Access Class 13 is configured in the UE.
14	Access Class 14 is configured in the UE.
15	Access Class 15 is configured in the UE.

Access Category is an indication of the type of access attempt by UEs that should be barred. Strictly speaking this is related to the service or application and not to the UE as such. However, it is taken into account that certain types of UEs can be dedicated to specific services/applications, such as terminals dedicated to emergency communications. Table 3 which is also included for illustrative purposes, provides an overview of the access categories:

Table 3. Access Categories used in 5G unified access control (extract from Table 6.22.2-1 in TS 22.261 [13])

Access Category number	**Conditions related to UE**	**Type of access attempt**
0	All	MO signalling resulting from paging
1	UE is configured for delay tolerant service and subject to access control for Access Category 1, which is judged based on relation of UE's HPLMN and the selected PLMN.	All except for Emergency
2	All	Emergency
3	All except for the conditions in Access Category 1.	MO signalling resulting from other than paging
4	All except for the conditions in Access Category 1.	MMTEL voice
5	All except for the conditions in Access Category 1.	MMTEL video
6	All except for the conditions in Access Category 1.	SMS
7	All except for the conditions in Access Category 1.	MO data that do not belong to any other Access Categories
8-31		Reserved standardized Access Categories
32-63	All	Based on operator classification

Based on this table, the UE can determine the type of access it may want to perform, and which types may be barred. E.g. a normal MO call may be barred, but not an Emergency call.

2.3 Rationale and evolution of net neutrality

The background of net neutrality and the related policies and regulations can be found in many excellent academic texts (see e.g. [21] and references therein). Two recurring high-level goals in net neutrality policies are:

- ensuring the freedom of end users in their choice of applications and services on the internet;
- ensuring continued innovation in applications in the internet ecosystem.

In the promotion and safeguarding of these goals, net neutrality policies focus at the role that Internet Service Providers (ISPs) have in the internet ecosystem through the Internet Access Services and the other services and applications they provide over their networks. Through their position in between Content and Application Providers (CAPs) and end users[4], ISPs could act as gatekeepers by introducing technical or commercial limitations on the traffic flow to/from end users and CAPs,

[4] In this introduction, end users are viewed as the users of services and applications offered by CAPs. The Regulation and Guidelines take a more refined view in which CAPs can also be end users.

potentially to the benefit of the applications they offer themselves. Such limitations have been reported in surveys (e.g., [22]) and have also appeared in public disputes between companies and in public debates (see [21] for examples). It is worth to note that companies other than ISPs can also steer or limit the access and use of applications and content on the internet by end users ([23], [24]). This typically involves actions in parts of the internet ecosystem that are outside the internet access segment and therefore outside the scope of net neutrality regulation.

Net neutrality is an area with a tradition of strong debate among stakeholders, in the Netherlands, in Europe, the US and in other regions. The debate includes many perspectives, from technical and business to societal and ideological. This mix of perspectives explains the complexity of the debate. It is challenging to weigh and reconcile the diverse and often rather fundamental interests of different stakeholders. Traffic management by ISPs and the relation to application requirements have been a central ingredient in the net neutrality debate from the very start. As observed by Tim Wu in his famous 2003 paper [25] in which he coined the term net(work) neutrality:

> "… IP was only neutral among data applications. Internet networks tend to favour, as a class, applications insensitive to latency (delay) or jitter (signal distortion)… . In a universe of applications, that includes both latency-sensitive and insensitive applications, it is difficult to regard the IP suite as truly neutral as among all applications."

Wu's observation is relevant again for 5G networks that aim to provide connectivity tailored to the requirements of specific applications. As can be expected, considerations on traffic management measures by ISPs that touch on Wu's observation are found in the policies and legal instruments introduced over the years in Europe and the US.

In the US, the FCC made its first statements on net neutrality in 2004 [26] while the first specific rules on net neutrality were introduced its 2010 Report and Order [27]. These rules allowed for "reasonable network management", where "reasonable" was to be interpreted as "appropriate and tailored to achieving a legitimate network management purpose, taking into account the particular network architecture and technology". The FCC published its latest set of net neutrality rules in its 2015 Open Internet Order [28]. These include three so-called "bright line rules": no blocking, no throttling and no paid prioritization by ISPs. In December 2017 though, the new FCC leadership voted to withdraw the rules altogether [29]. In California and several other US states, lawmakers are now considering introducing the rules at state level [30].

In Europe, the EC introduced several policy objectives in the area of net neutrality in 2009 [31]. The first rules introduced at the European level are contained in the 2009 universal service directive [32]. They focussed on introducing transparency, in particular in the area of traffic management by ISPs. The purpose of this transparency was to give end users a meaningful insight into the traffic management methods which are employed by ISPs and what consequences they have for them.

In parallel to the policies and rules introduced at the European level, the Dutch parliament introduced stronger national rules that moved beyond transparency and explicitly prohibited blocking and throttling (with some specific exceptions) [33][5]. The Dutch rules also explicitly prohibited price discrimination by ISPs based on the type of applications carried over the IAS. The best-known example of such discrimination is zero-rating, a commercial arrangement in which the data for specific applications or application groups does not count towards a customer's monthly data bundle.

In 2015, the EC introduced the current European rules for net neutrality in its Regulation 2015/2120 [2]. As discussed in the next section, the Regulation includes rules on traffic management which, among other things, explicitly prohibit blocking and throttling, but also provides several exceptions. The Regulation also provides rules that apply to commercial practices like zero-rating. Since the rules are contained in an EU Regulation (and not in a directive), they apply "as is" in all EU Member States[6]. The Regulation is accompanied by implementation Guidelines [3] developed by BEREC, the Body of European Regulators for Electronic Communications. The Guidelines have been developed for National Regulatory Authorities (NRAs) that have the task to enforce the Regulation.

2.4 Key points in EU Regulation and BEREC Guidelines

The regulatory analysis of 5G architecture options in this study is based on the Regulation and the Guidelines. The key points needed for the analysis are introduced below. Rules on commercial practices, such those that apply to zero-rating, are not covered here as they are outside the scope of this study. Transparency measures are also not considered, as these are unlikely to introduce limitations for the use of 5G architecture options. For more detail and context of the points below, the reader is referred to the Regulation itself and the Guidelines that also provide further explanation.

2.4.1 Starting point: description of the internet

An important starting point is the definition of the internet. In paragraph 14 of the Guidelines, BEREC understands (rather than defines) internet as "*a global system of interconnected networks that enables connected end-users to connect to one another*". Note that this description does not include terms related to particular technologies used, such as IP (Internet Protocol) or IP addresses. This makes the definition technology neutral. It therefore also applies to Ethernet connectivity[7].

Paragraph 4 of the Guidelines further describes the term "*end users*": it encompasses individuals and businesses, including consumers as well as Content and Application Providers (CAPs). This underlines that the rules introduced for the IAS (see next section) apply not only to the classical net neutrality setting where a consumer uses an IAS for access to services on the internet, but also when CAPs use an IAS in their provision of services over the internet.

[5] Slovenia introduced stronger rules as well.
[6] This is illustrated by a Dutch court decision that the Dutch zero-rating rules are not compliant with the Regulation. The Dutch telecoms law will now be changed to align it to the Regulation.
[7] One could argue that the description also covers the classical circuit-switched PSTN (Public Switched Telephony Network).

2.4.2 Definition and rules for Internet Access Service

The Internet Access Service (IAS) is defined in Article 2 of the Regulation: "*'internet access service' means a publicly available electronic communications service that provides access to the internet, and thereby connectivity to virtually all end points of the internet, irrespective of the network technology and terminal equipment used*."

Paragraph 10 of the Guidelines provides the interpretation of the term "*publicly available*". Services that are offered not only to a predetermined group of end users but in principle to any customer who wants to subscribe to the service are considered to be publicly available. If a service is offered only to a predetermined group of end users, then it is considered to be not publicly available. In practice, such offers are often characterised as "private". The Guidelines only use the term "private" in the context of networks.

Internet interconnections (through peering and transit agreements) are outside the scope of the rules for IAS, as they are not a part of the internet access. Still, they are mentioned in this context in paragraph 6 of the Guidelines: "*NRAs may take into account the interconnection policies and practices of ISPs in so far as they have the effect of limiting the exercise of end-user rights …*".

The crucial rules for traffic management in IAS are built up in three steps in Article 3.3 of the Regulation. The first step is an overall rule that states that all traffic shall be treated equally:

> "Providers of internet access services shall treat all traffic equally, when providing internet access services, without discrimination, restriction or interference, and irrespective of the sender and receiver, the content accessed or distributed, the applications or services used or provided, or the terminal equipment used."

In step two, an exception is made for so-called "*reasonable traffic management measures*" which is linked to a number of conditions:

> "*The first subparagraph shall not prevent providers of internet access services from implementing reasonable traffic management measures. In order to be deemed to be reasonable, such measures shall be transparent, non-discriminatory and proportionate, and shall not be based on commercial considerations but on objectively different technical quality of service requirements of specific categories of traffic. Such measures shall not monitor the specific content and shall not be maintained for longer than necessary."*

As seen later in this report, these conditions are important when assessing the use of QoS differentiation. In a third step, a further exception is made for more intrusive traffic measurement measures such as blocking and throttling. This further exception is linked to very specific situations. For the purpose of this study, the situation of "*exceptional or temporary network congestion*" is relevant.

> "*Providers of internet access services shall not engage in traffic management measures going beyond those set out in the second subparagraph, and in particular shall not block, slow down, alter, restrict, interfere with, degrade or discriminate between specific content, applications or services, or specific*

categories thereof, except as necessary, and only for as long as necessary, in order to:

(a) (..)
(b) (..)
(c) prevent impending network congestion and mitigate the effects of exceptional or temporary network congestion, provided that equivalent categories of traffic are treated equally."

2.4.3 Sub-internet Services and limited number of reachable end points

The definition of the IAS contains the phrase "*connectivity to virtually all end points of the internet*". In paragraph 17 of the Guidelines, it is made very clear that so-called *sub-internet services* that offer access to only part of the internet end points (e.g., only to selected websites, or the whole of the internet except for certain applications or services) are not allowed. According to the Guidelines, the fact that a service does not provide connectivity to (virtually) all end points does not imply that the rules for IAS do not apply. Rather, sub-internet services are viewed as IASs that fail to comply with the rules.

An exception is made in paragraph 18 for services where the number of reachable end-points is limited by the nature of the terminal equipment, rather than by the network operator that provides the connectivity. The examples mentioned are e-book readers and machine-to-machine devices like smart meters, where the connectivity service can be designed for communication with individual devices rather than with arbitrary end points on the internet. The guidelines offer room for such connectivity services, which can be viewed as a particular class of sub-internet services unless they are used to circumvent the regulation.

2.4.4 Specialised services

In the Regulation, specialised services only appear as "*other services*". The term specialised service is introduced in the BEREC guidelines. The room for network operators and CAPs to offer specialised services (SpSs) is described in Article 3.5 of the Regulation:

"*Providers of electronic communications to the public, including providers of internet access services, and providers of content, applications and services shall be free to offer services other than internet access services which are optimised for specific content, applications or services, or a combination thereof, where the optimisation is necessary in order to meet requirements of the content, applications or services for a specific level of quality*."

The "*other services*" terminology underlines that the Regulation introduces a binary split between the IAS and the SpS, i.e., there is not a third category. In paragraph 101 in the Guidelines, the *characteristics* of SpSs are grouped in three parts as:

- *they are services other than IAS services;*
- *they are optimised for specific content, applications or services, or a combination thereof;*
- *the optimisation is objectively necessary in order to meet requirements for a specific level of quality.*

For the purpose of this study, two further points on SpS characteristics in the Guidelines are important:

- *Paragraph 110 provides guidance for the assessment of the optimisation: "If assurance of a specific level of quality is objectively necessary, this cannot be provided by simply granting general priority over comparable content. Specialised services do not provide connectivity to the internet and they can be offered, for example, through a connection that is logically separated from the traffic of the IAS in order to assure these levels of quality."*
- *Paragraph 115 further addresses connectivity to the internet from SpSs in the context of Virtual Private Networks (VPNs): "VPNs could qualify as specialised services in accordance with Article 3(5) of the Regulation. However, in accordance with Recital 17, to the extent that corporate services such as VPNs also provide access to the internet, the provision of such access to the internet by a provider of electronic communications to the public should comply with Article 3(1) to (4) of the Regulation."*

The regulation also specifies *conditions* under which SpS may be offered in parallel to the IAS:

> *"Providers of electronic communications to the public, including providers of internet access services, may offer or facilitate such services only if the network capacity is sufficient to provide them in addition to any internet access services provided. Such services shall not be usable or offered as a replacement for internet access services, and shall not be to the detriment of the availability or general quality of internet access services for end-users."*

This is presented in paragraph 102 of the Guidelines as:

- *the network capacity is sufficient to provide the specialised service in addition to any IAS provided;*
- *specialised services are not usable or offered as a replacement for IAS;*
- *specialised services are not to the detriment of the availability or general quality of the IAS for end-users.*

3 Use cases to be supported by 5G

3.1 Three sector-specific use cases

The introduction of three sector-specific use cases in this study was motivated by the fact that the net neutrality analysis would otherwise remain too abstract. Working with use cases also automatically triggers the question which 5G functionalities should be taken on board and why. Conversely, in the selection of use cases, we have ensured that specific 5G functionalities are preferable or even conditional for their implementation. Three use cases can obviously never fully represent the much wider portfolio of use cases and application that will emerge over time. This is not a real problem as most of the 5G functionalities (also through consideration of different implementation options per use case), with relevance to net neutrality, are covered through the selection being made. The instrumental aim of the use cases is to extract from them a set of archetype architecture options which can be subjected to a net neutrality analysis.

The use cases relate to three distinct sectors: Media, ITS and Public Safety. These sectors (verticals) have been selected based on recognition of their relevance by the industry and telecom operators, their potential for special network treatment and on the knowledge position at TNO with respect to specific verticals. The choice for Public Safety was not without dispute as the arguments were put forward that firstly this sector is not clearly recognized as one of the prominent sectors in the 5G concept development and secondly that this institutional sector could probably qualify for a net neutrality waiver anyway. These were valid points, but on the other hand, the Public Safety sector in our view represents a wider group of *critical communication users* including non-governmental organisations and companies, in need of reliable connectivity and very high network uptimes, and for whom such a waiver would not apply. The need for ultra-reliable communications, being one of the three pillars of 5G, justifies this choice.

In the search for suitable use cases, the following criteria have been applied to the individual use cases:

- sufficiently specific for translation to network requirements and high-level implementation options;
- recognised in the industry (sector verticals and telecoms) as relevant for initial 5G deployment;
- sufficient background and industry knowledge available at TNO to make an independent assessment of application and network requirements.

Then, the three use cases together should:

- cover the key technical ingredients of the 5G architecture;
- lead to the identification of different potential issues and grey areas in the interpretation of net neutrality rules, such that it is plausible that a similar analysis applied to other use cases in the future would lead us back to the same issues.

3.2 Virtual Reality in media and entertainment

3.2.1 Industry context

In recent years, Virtual Reality (VR) has (re)gained substantial interest from researchers, large tech companies and start-up entrepreneurs. This is visible in the consumer market through the introduction and initial adoption of VR Head-Mounted Devices (HMDs) like the Samsung Gear VR, Oculus Rift and HTC VIVE. The appearance of the HMDs is accompanied by developments in VR content creation and distribution:

- Media broadcasters such as Sky [34], BT [35] and PCCW Singapore [36] have started their 360 VR broadcasts of sports events. The technology behind these broadcast is provided by start-up companies like NextVR [37].
- Facebook is adding VR to its social networking suite through its Facebook Spaces application ([38], [39], [40]). Facebook Spaces is positioned as a so-called Social VR application with the claim and catchphrase “VR is better with friends”.
- Amazon is also reported to move into VR, either as a part of their media offering [41] or a way to improve the user experience in its web shop [42].

VR in Media has been selected as a use case for this project as VR introduces several demanding network requirements. 5G networks will need to support VR applications as users are likely to expect VR applications to be available independent of their location and network connectivity. Note that the 5G specifications include WiFi connectivity. They also integrate the use of fixed access (such DSL, fibre and coax) in the overall 5G architecture.

Apart from the network requirements, VR is an interesting use case as it can lead to large volumes of streaming video. Today, internet traffic is dominated by video and the dominance is expected to grow further: from 67% of the total internet traffic in 2016 to about 80% in 2021 [43]. The large traffic volumes associated with video, combined with content provider and ISP strategies to handle those volumes, have earlier been an important driver of net neutrality debates [23]. Mass market adoption of VR would further add to this driver.

It is worth noting in addition to VR, there are related concepts such as Augmented Reality (AR) and Mixed Reality (MR) that have their own position on the Milgram scale [44], with the real world on one end and virtual reality on the other end. Microsoft seems to be focussing on MR and AR, for example through its HoloLens that can be used to superimpose virtual overlays on the physical world [45].

Obviously, there are many other application domains for VR, AR and MR outside media and entertainment. In the health domain, a very demanding application that is often mentioned is remote surgery. Unlike the media applications above, remote surgery is expected to occur only across well-controlled fixed locations (i.e., university hospitals) with dedicated fibre connections because of the stringent requirements on reliability. It is not expected that remote surgery will occur over the generic 5G infrastructure in the initial phases of 5G rollout.

3.2.2 360 VR streaming and social communication

The basic application in the VR in Media use case is 360 VR streaming of live or pre-recorded video. The video is recorded with dedicated 360-degree cameras and distributed over a network infrastructure to viewers with a suitable HMD.

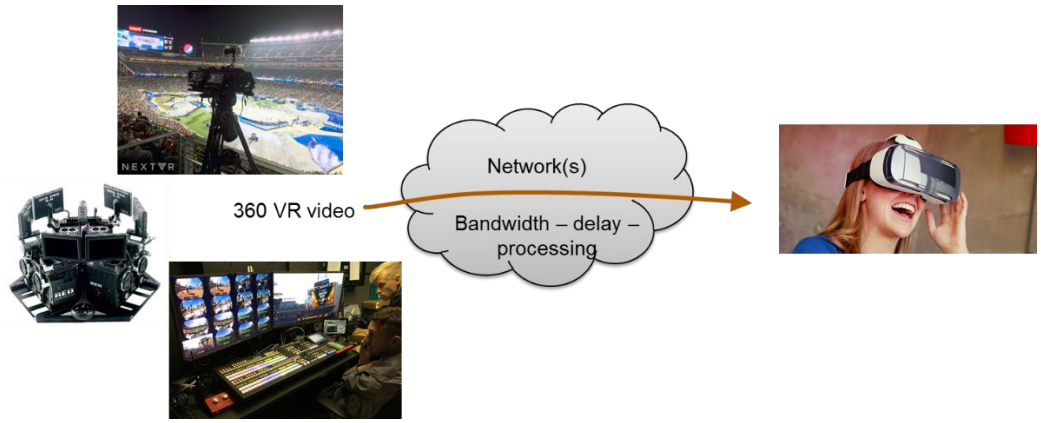

Figure 14: Streaming of 360 VR content to an HMD.

In social VR applications, the 360-degree content may not be limited to a single pre-recorded or live content stream but can also include user-generated and user-recorded content streams. These can then be combined into one integrated virtual image for the participating viewers. For the users, this has the added attraction that they can enjoy the content or event together and interact in a natural way.
The social VR application introduces additional stringent requirements, as multiple video streams need to be combined in real time. In particular, this introduces requirements for (image) processing in the network as the viewers' devices cannot be expected to have sufficient resources to do this. This is illustrated by the avatars that Facebook's social VR application currently uses as a surrogate for actual video images of the participants. In research settings, it has been demonstrated how the HMD can be removed from the recorded images of the viewers to make the image more natural in social VR settings [46].

3.2.3 Key application components and connectivity needs

Below, we analyse the data available in public literature on bandwidth and latency required for VR. These requirements apply to 5G networks used to support VR but also to other networks used for this purpose, as the requirements are determined by the application and not by the network. The requirements for bandwidth, latency and processing are interdependent and to a certain degree interchangeable. A common example of this, described in more detail below, is that a lower bandwidth can be compensated with specific technologies at the cost of more stringent requirements for latency. Thus, the content and network providers generally try to optimise the balance between application quality and requirements in a latency-bandwidth-processing triangle, see Figure 15.

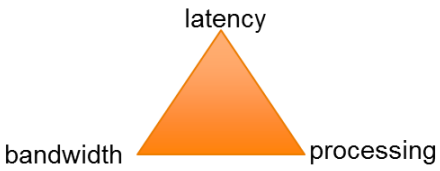

Figure 15: The VR requirements for latency, bandwidth and processing can be positioned in a triangle

As a starting point for the bandwidth in the triangle, this study assumes a requirement of the order of 100 Mbit/s. This is based on the requirements of current VR streaming applications, combined with a bandwidth increase that is needed for higher video resolutions. Today, a 360 video in 4K requires around 30-50 Mbit/s. It should be noted that the 4K resolution, which gives a high-resolution image on today's televisions, is used to cover the entire 360 sphere and gives only a modest resolution for the image segment actually visible on the HMD for viewers. Higher resolutions are therefore desirable. The 100 Mbit/s bandwidth used in this study is higher, but not excessively higher, than the highest bandwidth offered by well-known online video providers today: Netflix Ultra HD at 25 Mbit/s [47] and YouTube 4K/2160pixels @30 frames per second at 20 to 50 Mbit/s [48].

Adaptive viewport streaming and tiling techniques ([49], [50]) relax today's bandwidth requirement from 50 Mbit/s to around 25 Mbit/s or even to around 10 Mbit/s. This comes at the cost of tighter requirements for latency as a new area in the VR360 video needs to be shown whenever the user turns his head. The requirement for the so-called motion-to-photon delay is generally understood to be in the range of 20 to 40 ms ([51], [52], [53], [54]), with some sources mentioning smaller latencies on the order of 5 ms [55]. This includes the time needed for the detection of head movement and for server side to send a new viewport. Some of today's implementations use 100 ms as an acceptable latency, but this is essentially a suboptimal approach that aims at the latency achievable over the Internet today rather than at the actual requirement from the application. In VR applications that also include real-time streams from participating users, the low-latency requirements also apply to the image processing needed to stitch the different streams together and to remove undesired elements like the HMDs from the participant's images.

In summary, this study assumes a combination of 100 Mbit/s of bandwidth and network latency of 20-40ms, see Figure 16. These are not considered to be the most challenging requirements that could be determined for VR cases, but rather to be realistic requirements, which do introduce a significant challenge for networks.

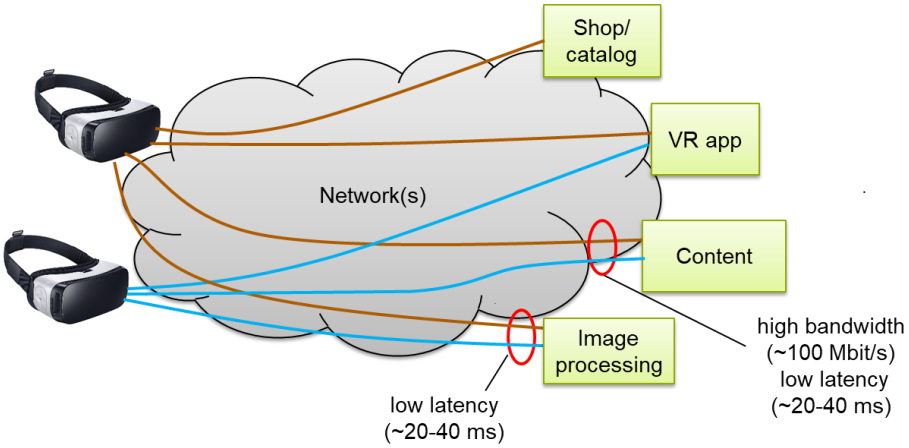

Figure 16: The 360 VR content streams and the user-generated content introduce high bandwidth and low-latency requirements.

It is also relevant to note that not all traffic involved in VR applications introduces demanding requirements. For example, the e-commerce interactions, in which a user purchases an app or content, and the downloading of the app do not introduce requirements that exceed the performance of the current best-effort internet. For the processing, we do not formulate quantitative requirements for the CPU, storage or other metrics. For the analysis in this study, the location of the processing (e.g., in the form of edge computing) is more important than the performance itself.

3.3 Critical communications in Public Safety

3.3.1 Industry context

Public Safety or Public Protection refers to typical activities of police, ambulance and fire brigade, focussed on maintaining public order and restoring situations back to normal in case of planned (large) events or unplanned incidents and disasters. In international literature, the term Public Protection and Disaster Relief (PPDR) is used, which explicitly addresses the deployment of emergency services also during specific calamities or catastrophic events.

Mobile communication means are essential to this group to carry out their tasks effectively. Due to the very specific functional and non-functional requirements in this sector which can be summarized in the term *mission critical communications*, dedicated technologies and solutions emerged in the past. A widely used and worldwide recognised standard is TETRA, which was developed in the nineties under ETSI authority [56]. TETRA standardises Professional Mobile Radio (PMR) and Public Access Mobile Radio (PAMR) technology which can deliver speech and short data services. P(A)MR technology offers a number of very specific features which, especially at the time of their development, could not be found elsewhere. The TETRA standard addresses solutions for a wider category of critical communications users: PPDR entities but also, for example, users in industry, airports and harbours. The custodianship of the TETRA technology roadmap and its deployment lies with the TETRA Critical Communications Association (TCCA). National networks for PPDR are often TETRA based and tailored to specific needs.

Under pressure of standardisation developments in the commercial cellular world, the TCCA changed its course in 2011-2012 when in the US LTE technology was considered a serious candidate for a new nationwide PPDR network (FirstNet [57]). The Critical Communication Broadband Group (CCBG) was erected and developments at 3GPP were endorsed and strengthened by bringing in sector-specific expertise into this standards development process. This led to subsequent 3GPP Releases – starting with Release 12 - with PPDR specific features in mission critical communications. PPDR professionals in the Netherlands and elsewhere are already subscribing to regular commercial mobile communication services for non mission critical data communications. The CCBG represents a wider group of Critical Communication users like Industry, Harbours, Airports, etc. The relevance of the PPDR use case in this report also extends to these user categories.

EU Member states find themselves in different situations and prospects, but the need for harmonisation in the procurement of Mobile Broadband solutions is recognized. This has very recently led to the launch of instruments to stimulate this harmonisation (i.e. PCP: Pre-Competitive Procurement and PPI: Public Procurement of Innovation Solutions). The Broadmap project [58] is the first of its kind which specifically targets European PPDR requirements and a high-level architecture such innovations should adhere to. This must be seen as an add-on to what is already standardized in 3GPP and helps to create a future European market for PPDR communication solutions which are based on generic commercial technology.

3.3.2 Applications

There is an increasing dependence on high-quality, secure and timely information in the operations of emergency services [59]. The network-centric working concept, with a much more horizontal rather than vertical (hierarchical) tasking, is gradually adopted. Information applications such as access to various databases, distribution of situational awareness information, uploading video/imagery, aggregation of sensor data, exchange of medical data, access to 3D maps, exchange of location and status data, have become a crucial part of the day-to-day work of PPDR workers. New technologies and instruments, like the use of sensors and drones, are being adopted and bring their own communication needs. New working concepts like remote medical support are arriving. The typical pattern in the adoption of new (information) applications is to try out, experience, assess, conduct practical trials, upscale and then adopt in working standards. Each new application begins as a small-scale experiment and could end up as one of the operational applications, with adopted working procedures. Mobile data growth in this domain, which currently lags behind, is expected to at least follow the societal trend which is predicted to be between 120 and 140% (CAGR) in Western Europe [60].

The portfolio of proven PPDR applications as such is not the most interesting in the context of this study. The key point of all PPDR applications is that they must be available and useable by authorized personnel *at all times, under all conditions.* This is what mission critical communications actually require, following the definition of the term *mission critical* as proposed by the Law Enforcement Working Party [61]. This rather holistic application requirement immediately translates into demanding mobile networking requirements which we will address in the next paragraph.

3.3.3 Application components and connectivity needs

The relevant application components are servers which reside in a confined intranet environment or are reachable through the internet. The high-level application requirement translates into demanding network requirements as follows:

- high system and service availability over a very broad spectrum of calamity scenarios;
- nationwide mobile network coverage (outdoor & indoor);
- ability to accommodate both planned and unplanned peak capacity demands and QoS requirements on individual applications in confined and crowded areas (generating high background traffic load). This capacity requirement particularly applies to the uplink;
- ability to pre-empt existing (non-PPDR) connections to allow new PPDR connections, in case of congestion.

These requirements have consequences for the delivery mechanism of connectivity services. It requires robustness in the network and special treatment of traffic to protect against local congestion situations.

A coarse distinction could be made between PPDR dedicated applications which are hosted in a secured (private intranet) area due to security and availability requirements and applications which can be found on the public web and are accessible to anyone. Therefore, plain internet access is included in the PPDR needs. It could be argued that public web-based applications are not made more robust for any particular user. This seems to invalidate this requirement. However, the concern of the PPDR community is less on the availability of a particular server on the world wide web as other applications on the web might provide substitutional information. The concern is much more on internet access limitations due to load and congestion effects in the mobile network, especially during larger incidents. Normal users, in much greater numbers in the incident area, collectively claim a much larger capacity than under normal circumstances.

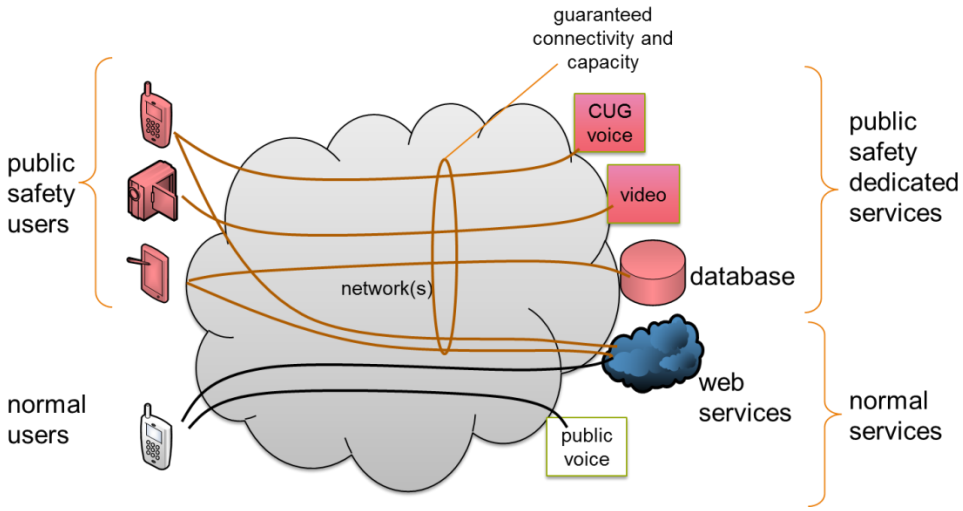

Figure 17: PPDR users using dedicated as well as generic applications (information, voice, video).

3.4 Automated Driving

3.4.1 Industry context

The development of Intelligent Transport Systems (ITS) began in the 1990s and went through a gradual process of concept definition, technology developments, standardisation and trials. Large-scale deployments are the logical next step in this sequence. In Cooperative ITS (C-ITS), the focus to date has been on two related tracks: (1) OEM (Original Equipment Manufacturer = Car Manufacturer) driven vehicle-to-vehicle communications (V2V) to allow vehicles to become aware of each other's proximity and to facilitate cooperative behaviour; (2) Traffic Management driven roadside to vehicle (and vice versa) information exchange (V2I), allowing roadside systems to gather traffic data, and allowing vehicles (and/or their drivers) to receive relevant information on road conditions, traffic lights, etc. Initially, these services were foreseen to be built on an ITS-specific protocol stack (ITS-G5), requiring dedicated communication facilities. This included a specific roadside infrastructure for the second track. However, with the nation-wide availability of high-grade generic mobile communication services (3G/4G), less time-critical ITS services implementations also became available via those channels in a cellular or so-called hybrid approach (as in projects like A58 Spookfiles and Talking Traffic which are part of the Dutch 'Beter Benutten' ITS programme [62]). The portfolio of ITS services and their corresponding message sets continue to be extended and harmonized in Europe (through projects like InterCor and C-ROADS [63][64]). A massive roll-out and adoption of ITS has however not yet taken place.

In the meantime, the transport and mobility ecosystem worldwide has entered an era of transformation, fuelled by technological factors (materials, electronics and sensors, pervasive internet, artificial intelligence) as well as societal factors (urbanisation, climate change, the shift towards sustainability, and to service and sharing models). After a long incubation period, ITS is now catalysed by important innovations in vehicles, i.e. electrification and automation of driving functions, also known under the terms automated and autonomous driving. It is also catalysed by

connected car and Over-The-Top application developments which originally consisted of an exchange of diagnostics data (M2M communications) and delivery of infotainment services. Today, the telecommunications industry has a much stronger ambition in serving the automotive domain. Evidence of this ambition is the recent establishment of the EATA (European Automotive and Telecom Alliance) [65] under endorsement of the EC, the adoption of specific connectivity requirements in standards for 5G and the execution of tests and trials with telecoms and automotive participation (in projects like 5G SAFE [66] and CONCORDA [67]).

It can be expected that Cooperative ITS, Connected Car-OTT approaches will converge, adopting the best of both worlds. The resulting outlook is a growing immersion of vehicles and their functions with the cloud, leveraging future internet and broadband telecommunications technologies and involving multiple players. An example of this is sketched in Figure 18.

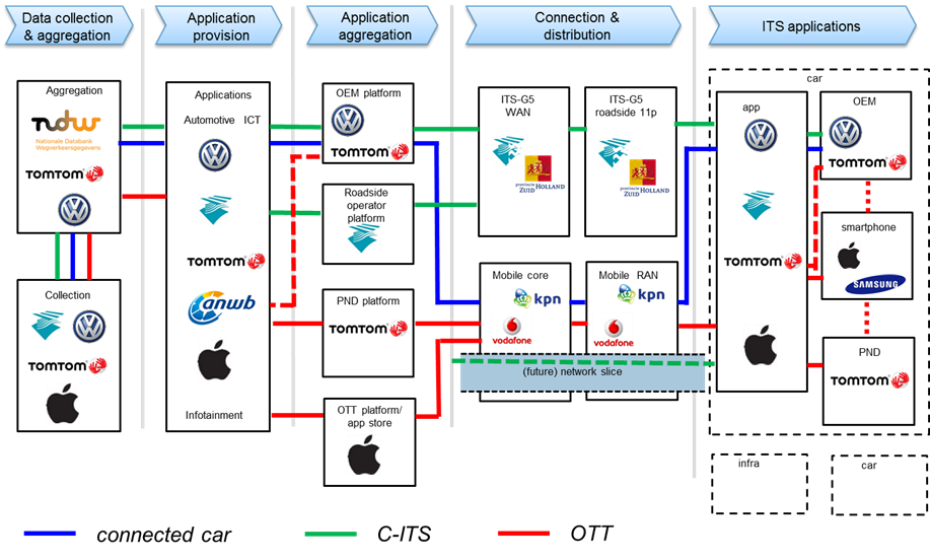

Figure 18: Ecosystem example for ITS, involving Connected Car, C-ITS and Over-The-Top services. Service chains and company logos are only for illustration. Source: TNO

3.4.2 Applications supporting automated driving

Through time, vehicles will maintain a certain required level of autonomy but also make use of sophisticated (private or public) cloud services. Examples of services supported by clouds are smart navigation and enhanced driving and manoeuvring functions using Local Dynamic Map data for increasing the vehicles' environmental perception beyond their own sensor range. In turn, vehicles will act as a data source themselves, contributing to the perception and influencing the behaviour of others. For various reasons, large-scale direct inter-vehicle communications will not be realizable any time soon, which increases the importance of cloud services to support situational awareness of vehicles. Cloud services can leverage data analytics and smart city type platforms via IoT and open data (as explored in European Horizon 2020 Large Scale Pilots [68]). Information services targeting the human driver will remain but change along with the changing role of the driver. Smart vehicles interacting electronically with other local and remote entities is a sophisticated form of Machine to Machine communication. As the driver in an

automated vehicle will get time available to do other things, entertainment services are expected to enter the vehicle, e.g. augmented reality projections on the windscreens and side windows.

The success of services and applications that support automated driving will depend on whether data can be offered to vehicles in a secure, reliable and timely fashion, with sufficient data speeds. The natural tendency particularly of traditional OEMs is to treat the automotive part of the vehicle as a closed and autonomously operating subsystem, i.e. with a complete set of onboard sensors and a minimum of interaction with external data sources. This concept is being challenged, as the use of external data services could reduce the materials bill considerably and could help introducing automation in vehicles in lower price segments. However, a breakthrough will only be possible if stringent requirements on security, reliability, timeliness and capacity can be met.

3.4.3 Key application components and connectivity needs

Table 4 shows example applications and their connectivity requirements published by 3GPP. The examples mainly address V2V connectivity but apply equally to V2N connectivity.

Table 4. ITS application categories and associated connectivity requirements. Source: 3GPP TS 22.186 V15.2.0 (2017-09)

Application	Degree of Automation	Max. Latency (ms)	Reliability (%)	Data rate (Mbit/s)	Min. Range (m)
Platooning (excl reporting)	Lower degree	10-25	90	Unspecified	350
	Higher/highest degree		99.99	65	80-180
Advanced Driving	Lower degree	25-100	90	10-50	700
	Higher degree	10-100	99.99-99.999	10-50	350
Extended Sensors	Lower degree	50-100	90-99	10	100-1000
	Higher degree	10-50	95-99.999%	10-1000	50-1000
Remote driving	Unspecified	5	99.999%	25 (UL) 1(DL)	Unspecified

Obviously, latency is a crucial parameter in ITS, because mobility is inherently characterized by scenario/situation dynamics (speeds and accelerations) while at the same time safety of life requirements apply. Deeper levels of automation mean that more time-critical automotive functions are automated with a certain dependency on external data in the process. Latency is a parameter of great influence in closed-loop control systems formed by automated vehicles. The use of external data brings latency of data communications into the vehicle's control loops (observe-analyse-decide-act). Latency values which go beyond the specification can create instabilities in such systems. Vehicles can and will apply time gating to prevent this effect but then the ignored (aged) data no longer positively contributes to the vehicle's automated performance. Hence, for the data to be useful in these control loops, it needs to adhere to the latency specifications. The maximum amount of latency that can be tolerated depends strongly on the system application considered. The 3GPP table above suggests that latency values (end to end) must be less than 100 ms, with a few applications that have more stringent requirements. This is confirmed by an extensive investigation into V2X communication requirements at the University of Aalborg in 2015 [69].

OEMs generally apply high system reliability standards. In case of automated driving, these standards would also need to be imposed on processes external to the vehicle. The way around this is to give external data a lower reliability classification (in terms of availability and integrity) than its onboard data sources in the decision-making process. However, this reduces the added value of external data to the automation process. The reliability requirement depends on the specific automotive function and the desired level of automation. In the context of this study with a focus on 5G connectivity services, the following definition of reliability is used[8]:

> *Reliability of connectivity service: Fraction of time that (for any user at any supported location) the connectivity service meets the minimum requirements regarding QoS (incl. latency, throughput, etc.) as specified in the SLA.*

A “four or five nines” for the reliability of the connectivity is likely to become the norm for fully automated driving.

In Figure 19 the high reliability/low latency requirement are assigned to the connections involving vehicles, vulnerable road users (equipped with smartphones), traffic light systems and traffic management and automated driving (AD) supporting services. These connections represent data exchange processes which have an immediate effect on vehicles behaviour and traffic safety. It is important to emphasize that this type of data exchange often has a very local nature and relevance. Other mobility (information) services depicted in the figure do not have such requirements.

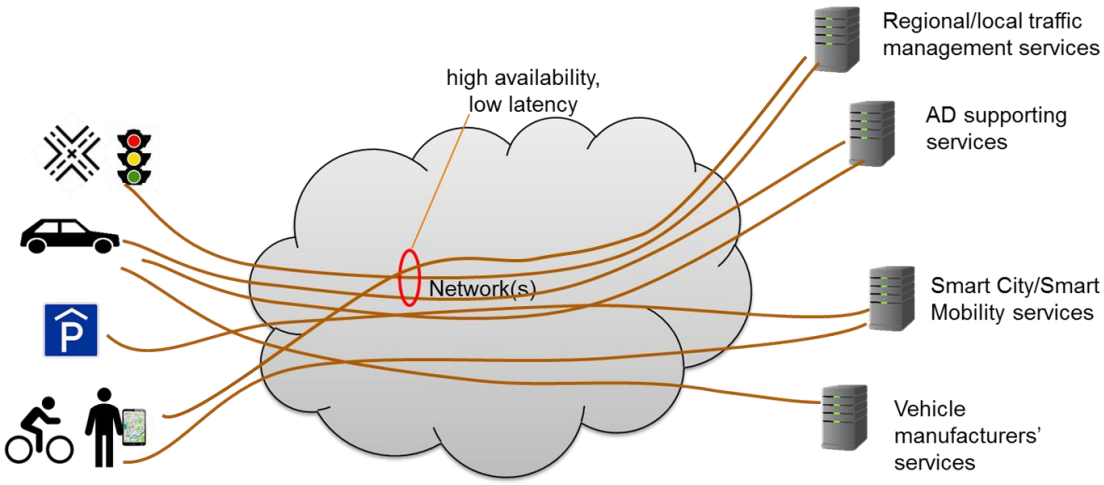

Figure 19: Future automated vehicles require high-performance interactions with (surrounding) services, involving low latency and high reliability.

[8] This definition has been proposed by Prof. Hans van den Berg, Twente University.

4 5G architecture options

4.1 Introduction to mapping of use cases to architecture ingredients

In this chapter, the use cases described in chapter 3 are mapped onto the 5G architecture ingredients as described in section 2.2. For each use case, multiple mappings are possible and have been investigated. In this chapter, only a selection of these mappings is elaborated, chosen in such a way that all the 5G features are encountered at least once. This provides an overview of the full range of applicability and use of these features.

4.2 VR in Media

For the virtual reality use case, four architecture mappings have been considered (Table 5).

Table 5. Four architecture mappings for the VR in Media use case. The mappings highlighted in blue are elaborated further.

Mapping	Slicing	Local Access	Service Hosting	QoS Differentiation	Access Control
VR#1		X			
VR#2		X		X	
VR#3		X	X		
VR#4	X	X	X		

The VR#1 mapping assumes the use of the 5G feature Local Access (for certain information streams) where this access would lead to the internet (i.e. outside of the operator domain). This use case mapping is further elaborated in section 4.2.1. and illustrated in Figure 20.

The VR#2 mapping assumes that in addition to Local Access also QoS Differentiation is applied for the (local) access to the internet.

The VR#3 mapping assumes the use of the 5G feature Local Access in combination with Service Hosting, i.e. the local access would lead to a service hosting environment inside the operator domain. Within this service hosting environment, operator services and 3rd party services can be accessed.

The VR#4 mapping assumes the use of the 5G feature of Slicing in combination with Local Access and Service Hosting. This use case mapping is further elaborated in section 4.2.2. and illustrated in Figure 21.

4.2.1 VR in Media based on Local access to Internet (VR#1)

In this mapping, the VR use case is mapped to the architecture options of Local Access *without* Service Hosting, i.e. for those data streams that need low latency or high bandwidth the mobile network will be exited near the user towards local content and local image processing. After exiting the mobile network, the data streams will leave the operator network and enter the internet. Note that in this case the internet is assumed to be extended to locations near the user.

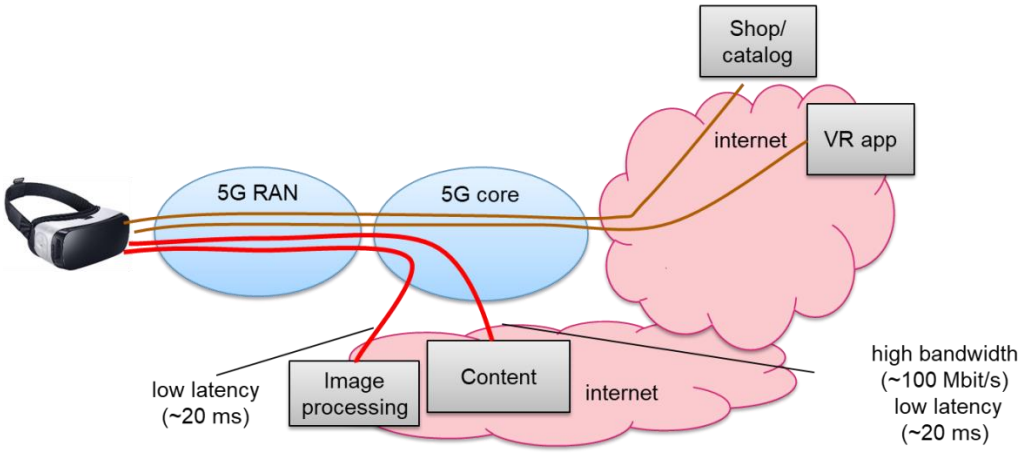

Figure 20: Local access architecture for the VR in Media use case, with the local access streams for image processing and content in red.

4.2.2 *VR in Media based on Slicing & Service hosting (VR#4)*

In this mapping, the VR use case is mapped to a combination of the architecture options of Network Slicing and Local Access *with* Service Hosting. Data streams that need low latency and/or high performance will make use of a specific VR slice *and* in this slice exit the mobile network near the user towards local content (which may need to be provisioned from the internet, as indicated by the dashed line) and local image processing. After exiting the mobile network in the VR slice the data streams will enter the operator network's Service Hosting Environment, where operator-controlled or trusted 3rd party-controlled applications operate.

Traffic that does not require special treatment is carried in another slice, which here is assumed to be an enhanced Mobile Broadband (eMBB) slice where the exit from the mobile network is to the standard internet at central locations.

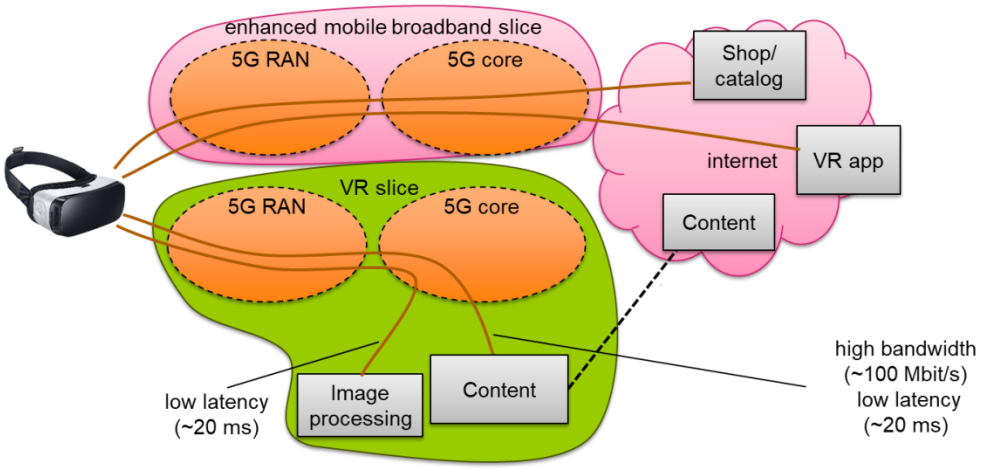

Figure 21: Slicing & service hosting architecture for the VR in Media use case.

4.3 Public safety communications

For the public safety use case, two architecture mapping have been considered as described in the following table.

Table 6. Two architecture mappings for the public safety use case.

Mapping	Slicing	Local Access	Service Hosting	QoS Differentiation	Access Control
PS#1				X	X
PS#2	X				

The PS#1 mapping assumes the use of the 5G features of QoS Differentiation and Access Control. Although these two features are in some form already existing in 4G networks, they are enhanced and improved in 5G networks. This use case mapping is further elaborated in section 4.3.1 and illustrated in Figure 22.

The PS#2 mapping assumes the use of the 5G feature of Slicing. This use case mapping is further elaborated in section 4.3.2 and illustrated in Figure 23.

4.3.1 Public safety communications based on Access control & QoS differentiation (PS#1)

In this mapping, the public safety use case is mapped to a combination of Access Control and QoS Differentiation. The use of (Unified) Access Control can enable access to the network for special groups (such as Public Safety workers) in case of extreme congestion situations (by barring other ‘normal’ users). Extreme situations are e.g. major disasters and other unplanned catastrophes. In less extreme, but still exceptional situations, the ARP part of QoS Differentiation can be used to enable capacity for special users by pre-empting (‘pushing out’) of normal users. For generic use, special QoS classes (5QI values) can enable better quality for special users, i.e. lower latency and less packet loss.

Public safety users may use the mobile network to access private public-safety-specific data networks. They may also use the mobile network for internet access, e.g. to obtain information available on the internet in support of their work during emergencies.

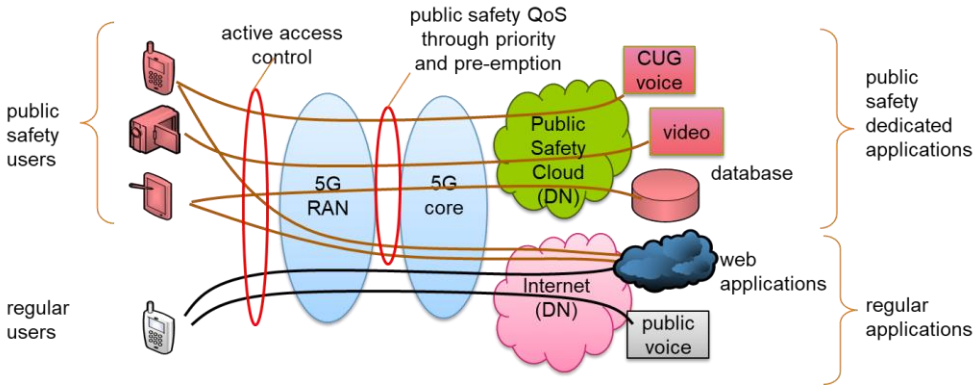

Figure 22: Access control & QoS differentiation architecture for public safety use case.

4.3.2 Public safety communications based on Slicing (PS#2)

In this mapping, the public safety use case is mapped to Network Slicing. The idea is that public safety users make use of a specific public safety slice and through this mechanism have a guaranteed capacity. This usage of slices assumes – as required by 3GPP – that slices can be assigned a 'capacity'. It is not clear yet if and how 3GPP will enable this assignment of capacity to slices. It is also not clear if slices can take capacity from other slices in emergency situations. In the non-normative 3GPP document TS28.801 [16], a use case is described where a slice that notices a need for capacity may take this from other slices. The corresponding normative specifications on 5G slice management have not yet been completed. Figure 23 assumes the capacity is predefined within the slice.

It should again be noted that the public safety slice provides access to both a private public safety data network and to the internet.

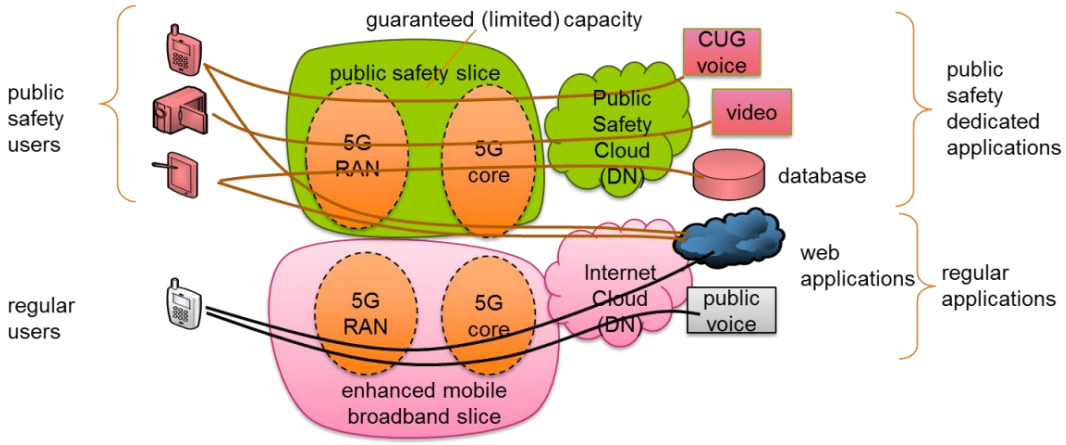

Figure 23: Slicing architecture for public safety use case.

4.4 Automated Driving

As described in the table below, four architecture mapping have been considered for the automated driving use case.

Mapping	Slicing	Local Access	Service Hosting	QoS Differentiation	Access Control
AD#1				X	
AD#2	X				
AD#3		X		X	
AD#4	X	X			

The AD#1 mapping assumes the use of the 5G feature QoS Differentiation (for certain information streams). Although this feature is in some form already existing in 4G networks, in 5G networks it has been enhanced and improved (especially in the area of extremely low latency).

The AD#2 mapping assumes the use of the 5G feature Slicing.

The AD#3 mapping assumes the use of the 5G features Local Access and QoS Differentiation. This use case mapping is further elaborated in section 4.4.1 and illustrated in Figure 24.

The AD#4 mapping assumes the use of the 5G features Slicing and Local Access. This use case mapping is further elaborated in section 4.4.2 and illustrated in Figure 25.

4.4.1 Automated Driving based on Local access & QoS differentiation (AD#3)

In this mapping, the automated driving use case is mapped to a combination of Local Access and QoS Differentiation. In this case, low latency and high availability are achieved without using special slices. Local access is enabled to a local network: the local ITS Cloud. However, the local network will also have connectivity (outside the operator network) to central networks: the central ITS Cloud. In addition to local access, QoS Differentiation is used to enhance the performance of the connectivity to the local network.

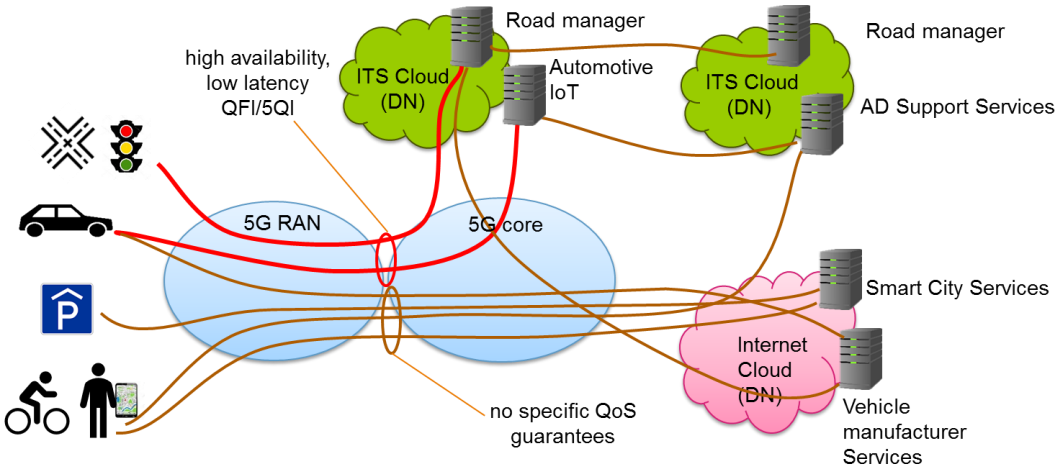

Figure 24: Local access & QoS differentiation architecture for public safety use case, with the local access streams to image processing and content in red.

4.4.2 Automated Driving based on Local access & Slicing (AD#4)

In this mapping, the automated driving use case is mapped to a combination of Local Access and Network Slicing. Local access to a local network (the local ITS Cloud) is enabled, but the local network will also have connectivity outside of the mobile network to central networks (the central ITS Cloud). Network Slicing is used to provide customized QoS. Within the special slice, QoS differentiation is arranged to allow a difference in treatment of the various traffic flows.

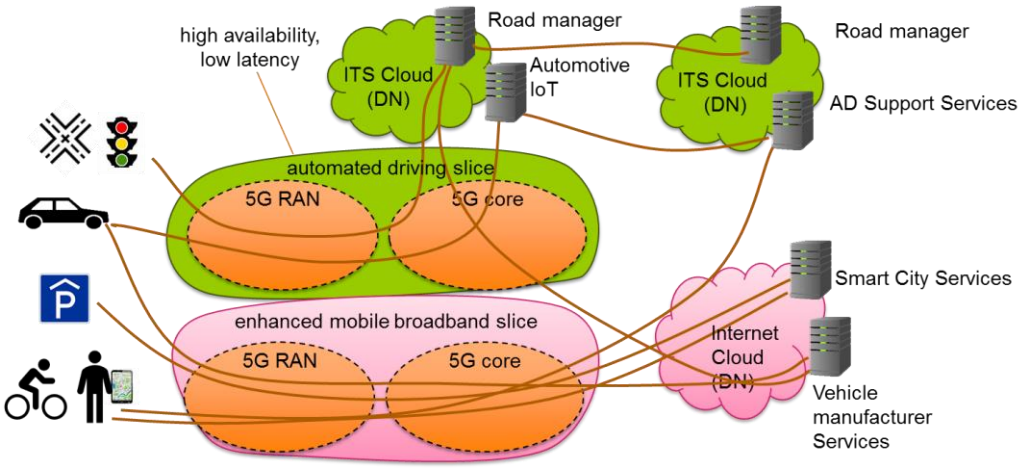

Figure 25: Local access and slicing architecture for public safety use case.

4.5 Consolidated 5G architecture view

For the evaluation of the alignment between the 5G architecture options and the net neutrality rules in the next chapter, it is convenient to have a consolidated view of the 5G architecture options. This view is presented in Figure 26. It abstracts from the individual options that have been developed in the different use cases. At the same time, it shows how archetype architecture options can co-exist in a single mobile operator network that has connectivity to the internet and to multiple other IP networks.

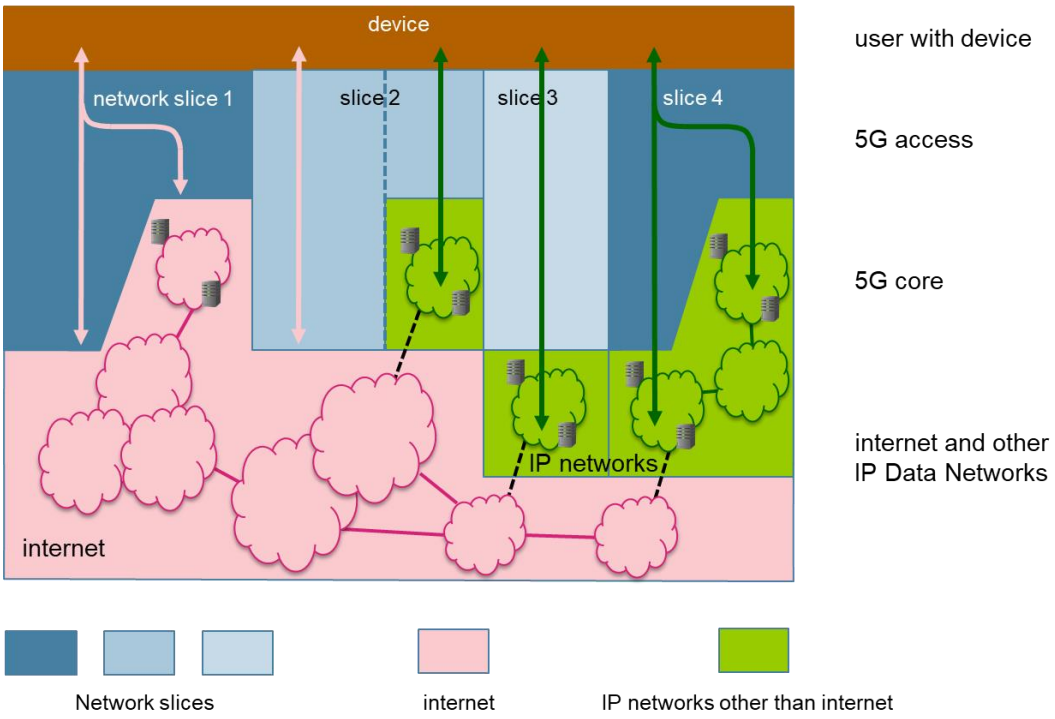

Figure 26. Consolidated 5G architecture view with multiple slices in a single mobile operator network.

The consolidated architecture model contains several parallel network slices. Some, but not all, slices provide connectivity to the internet. The internet is a prominent example of a Data Network (DN). The internet itself is made up of many interconnected IP networks (see section 2.4.1), with geographical scales that vary from global to national and regional. In the case of local access to the internet, one or more networks in the internet extend up to the geographical locations at the edges of the 5G mobile network. Apart from the internet, there are also other IP-based DNs, such as the ITS cloud in the autonomous driving use and the PS cloud in the public safety use case. Further examples of DNs are the IMS (IP Multimedia Subsystem) network used for VoLTE (Voice over LTE) in 4G networks, corporate VPNs and the IPTV platforms found in fixed networks.

As illustrated in Figure 27, the architecture options for the use cases presented in sections 4.2 to 4.4 can each be mapped to specific parts of the consolidated architecture view. Some of the options map to one slice, while others map to multiple slices.

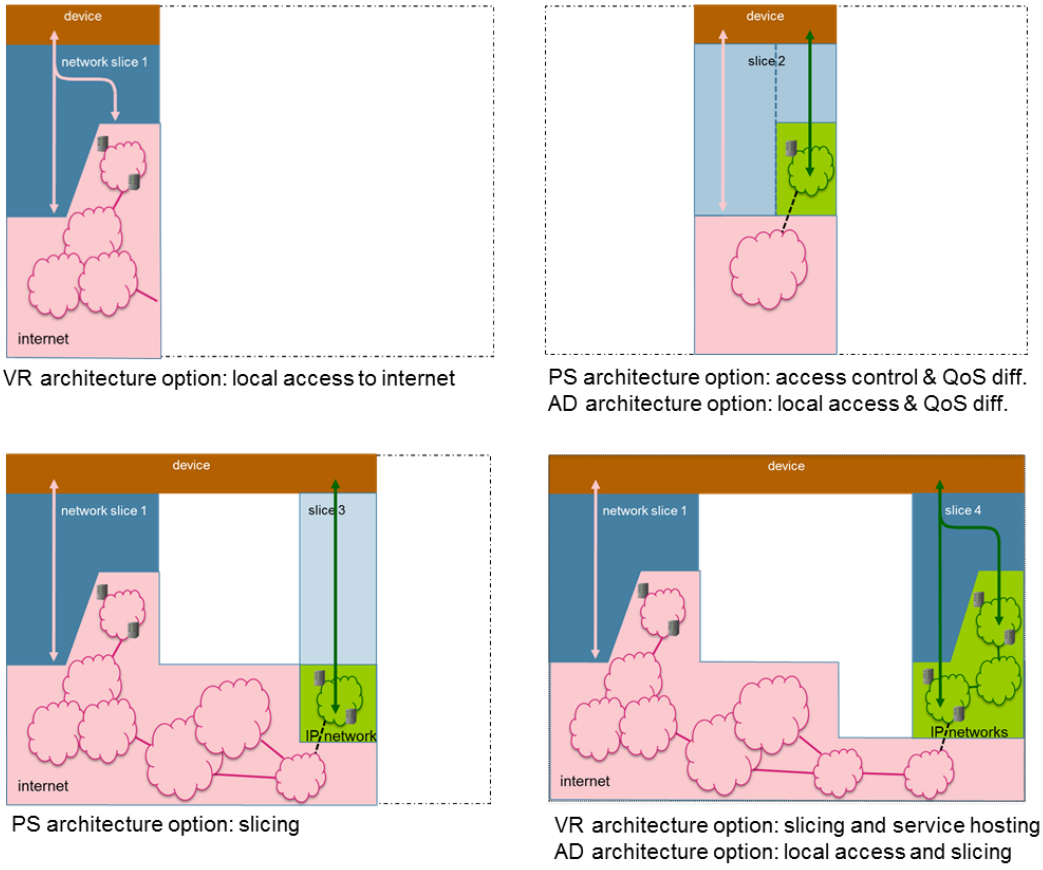

Figure 27: The individual architecture options for the use cases can each be mapped to the consolidated architecture view.

For the purpose of this study, we map architecture options from different use cases to the same slices. For example, in the bottom-right part of the figure, the VR architecture option “slicing and service hosting” and the AD architecture option “local access and slicing” both use slice 1 and slice 4. In practice, a network operator that provides support for both use cases could very well prefer to use separate slices instead of a common slice 4. This does not affect the analysis of the alignment between the 5G architecture options and the net neutrality rules in the next chapter.

5 Alignment of 5G architecture elements with Net Neutrality rules

5.1 Introduction

After the preparatory analyses in the previous chapters, we now arrive at the focal point of this study: the assessment of the alignment between the 5G mobile network technologies and the net neutrality rules. We start with a top-down determination of the parts in our consolidated architecture model where the rules for IASs and for SpS are expected to apply. We then zoom in on a number of individual points for a more detailed analysis of the alignment.

5.2 The distinction between net neutrality concepts and 5G architecture elements

Our top-down approach is inspired by the clear distinction in the Regulation between IASs and “other services”. The term “other services” underlines that the Regulation does not consider other categories than these two. In the remainder of our analysis, we will follow BEREC and use the term “Specialised Services” (SpS). With the binary split in the Regulation between IAS and SpS in mind, we can identify in which segments in our consolidated architecture the services, applications and traffic flows will need to comply with the rules for IAS. These are the segments where a publicly available service provides access to the internet, in accordance with the IAS definition from the Regulation (see section 2.4.2). In the remaining segments, the services will need to comply with SpS according to the logic of the binary split. The outcome of this identification is shown in Figure 28. Note that there is also a split between public and private services that can enter the analysis at another level. This is discussed further in section 5.7.

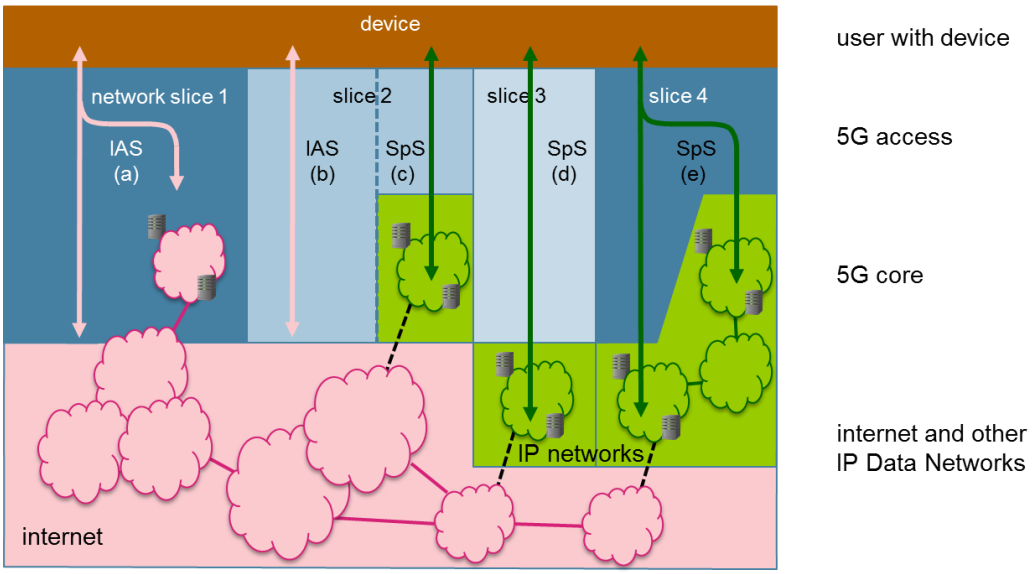

Figure 28: Identification of segments in consolidated architecture where services and applications need to comply with the rules for Internet Access Services (IAS) and the rules for Specialised Services (SpS).

In the next sections, we investigate whether the services, applications and traffic flows in different architecture options comply with the rules for IAS and SpS, as this is not self-evident from the top-down exercise. Nonetheless, the exercise already shows that it is crucial to distinguish between 5G architecture elements on the one hand and the net neutrality concepts of IAS and SpS on the other. Figure 28 clearly shows that there is no one-to-one mapping between the two. Two crucial examples of this are:

1. **A slice is not the same as a specialised service**. In 5G architectures that use slicing, an IAS is always in a slice. A slice can be used exclusively to provide an IAS (slice 1 in Figure 28). Alternatively, a single slice can be used to simultaneously provide an IAS and an SpS (slice 2). A slice can also be exclusively used to provide an SpS (slice 4).

2. **The application of QoS differentiation is not limited to IAS**. QoS differentiation can be used as a method for traffic management within an IAS. This could be done, for example, in slice 1 and 2 in the figure. However, it can also be used to assure the quality of SpSs (for example, also in slice 2). A prominent example of the latter is the VoLTE architecture in 4G networks, where specific QoS markings (the so-called Quality Class Indicators or QCIs) are used to prioritise and separate the VoLTE voice and related IMS signalling from the IAS traffic [70].

These two examples show that what matters for the assessment of compliance with net neutrality rules is how the 5G technologies are used, rather than the technologies themselves. This may seem like a trivial observation, but the history of net neutrality discussions shows that the basic concept of technological neutrality is easily forgotten.

5.3 Evaluation of the alignment of 5G architecture options with net neutrality rules

In the next sections, we assess in more detail whether the services, applications or traffic flows in different segments of the architecture comply with net neutrality rules. As noted earlier, the initial attribution of services and applications to either IAS or SpS does not necessarily imply that their implementation in 5G architectures is aligned with the rules. The next step is, therefore, to analyse in detail the mapping between the 5G architecture options and the net neutrality rules. This includes determining whether there are areas where the analysis cannot be completed because of multiple potential interpretations. Based on this analysis, we also give an indication of the relative complexity of each topic. For the purpose of this study, we define this as the relative complexity expected to be encountered by NRAs, network operators and CAPs in the analysis of specific cases that have more context information and (quantitative) details than the functional, use case-inspired analysis made here. Figure 29 shows the topics in the consolidated architecture that are analysed in more detail.

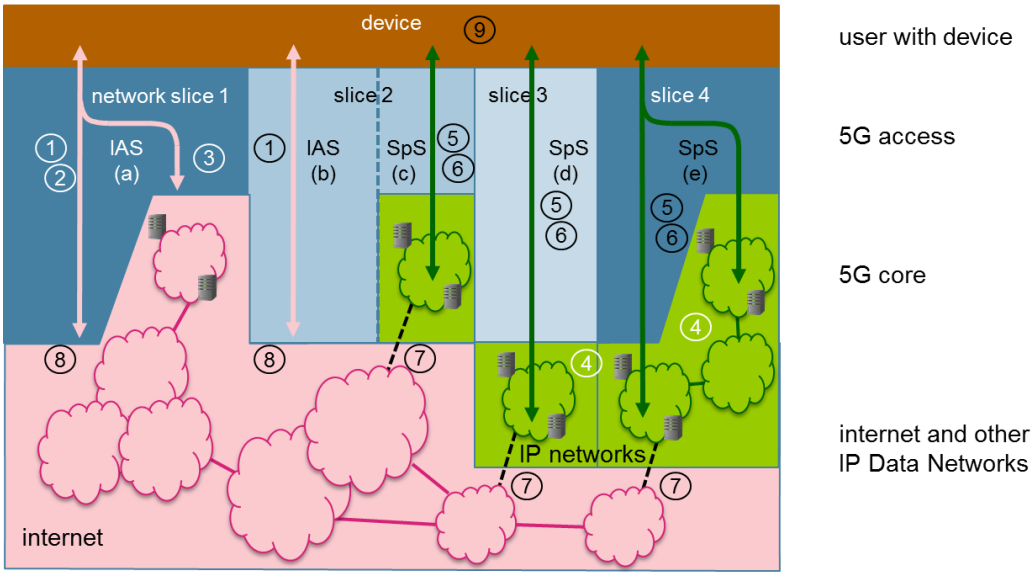

Figure 29: Topics in the consolidated architecture where the alignment between net neutrality rules and 5G architecture options is investigated in more detail.

The topics can be briefly described as:

1 Multiple IASs with different traffic management settings in one network
2 QoS differentiation within IAS
3 Local access to the internet
4 Public and private services and associated networks
5 Objective need for optimisation in SpS
6 Impact of SpS on IASs
7 SpS and connections to the internet
8 Connectivity to a limited number of internet end points
9 Access control

Before we go into the analysis, it is necessary to further characterise the device that is in the top part of the consolidated architecture. The devices encountered in our

use cases are diverse: they range from mobile phones and VR HMDs to embedded systems in cars. In this study, we are interested in uncovering all potential interactions between services, applications, traffic flows and networks slices that are relevant from the net neutrality perspective. We, therefore, assume that the user has a single device (UE) with one mobile subscription and identifier[9] where all connections come together. In practice, the use of multiple subscriptions, identifiers and devices in the provisioning of the services and applications can remove certain interactions. This can be useful in specific implementations but not in our analysis[10].

5.4 Topic 1: Multiple IASs with different traffic management settings in one network

The consolidated architecture model in Figure 29 contains multiple IASs that potentially have different technical characteristics that may originate in different traffic management settings applied by the mobile operator. For example, the characteristics of the IASs in slice 1 and in slice 2 can differ. Furthermore, within each of the slices there can also be multiple IASs with different characteristics (not separately shown in the figure). Here, it is important to recall that all services in the consolidated architecture are supported by a single mobile network operator. The offering of different IASs in one operator network is common today. A well-known example is provided by the separate consumer and business internet services that typically have different concentration factors in the aggregation network. In 5G networks, one can envision further differentiation, for example, IASs with traffic management tailored to IoT traffic. Another example could be the parallel offering of IASs with and without QoS differentiation.

When assessing the co-existence of multiple IASs in one network, the scope of Article 3.3 in the Regulation is important: *"providers … shall treat all traffic equally … irrespective of …**the sender and receiver**"* (bold added). The scope in which the sender and receiver are viewed determines the strictness of the rule and:

(a) Whether an ISP can provide different IASs to different end users and
(b) Whether an ISP can provide multiple different IASs to a given end user.

In the interviews held in the context of this study, we have found that the interpretation of Article 3.3 varies among experts. The majority interpretation is that the scope is an IAS provided to a given end user. This implies that:

a) One end user can have multiple IASs with different traffic management settings, e.g. one for general purpose internet access and one for autonomous driving applications with optimisations for IoT. These IASs can be in the same slice or in different slices.
b) Different end users can have different IASs with different traffic management settings, e.g. a consumer and a business service.

[9] This identifier can be the Subscriber Permanent Identifier (SUPI), which is the 5G successor of the International Mobile Subscriber Identity (IMSI). The SUPI is stored on a Universal Subscriber Identity Module (USIM), which can be a SIM Card.
[10] Note that the services in each of the slices in Figure 29 are in general provided to many end users that share the capacity in the slices. The assumption made here is thus that every user uses a single device for all connectivity.

The traffic management within each IAS needs to comply with Article 3.3.

Two alternative interpretations assume a wider scope for Article 3.3:

a) If the scope is assumed to be all the internet access traffic of one end user, then it is not allowed to offer different IASs with different traffic management settings aimed at supporting different application groups.
b) If the scope is taken to be all the IASs provided to all end users by a given ISP, then it implies that there can only be one type of traffic management setting applied across all IASs. This would prevent the co-existence of separate consumer and business IAS offerings.

In the remainder of this study, we assume that the intention of the Regulation is best reflected in the majority interpretation. We thus assume that it is allowed to have multiple IASs with different traffic management settings for a given end user.

Once the interpretation of *sender and receiver* has been clarified, the analysis of this topic is straightforward with low complexity.

5.5 Topic 2: QoS differentiation within IAS

The QoS differentiation mechanisms in 5G provide the technical basis for traffic management aimed at specific traffic flows within a user's traffic carried in an IAS. In the assessment whether such traffic management is allowed, the second step in the rules for traffic management in Article 3.3 is crucial (see section 2.4.2). This is the step that describes the conditions under which the traffic management is considered to be "reasonable" and therefore allowed under the Regulation:

> "*.. such measures shall be* ***transparent, non-discriminatory*** *and* ***proportionate****, and shall not be based on commercial considerations but on* ***objectively different technical quality of service requirements of specific categories of traffic****. Such measures* ***shall not monitor the specific content*** *and* ***shall not be maintained for longer than necessary.****"* (bold added)

Below, we make a number of observations on the interpretation of the conditions that jointly define "reasonable traffic management", based on the use cases from chapter 3 and the architecture options developed in chapter 4.

5.5.1 Objectively different technical quality of service requirements

The condition that the QoS differentiation is based on ***objectively different technical quality of service requirements of specific categories of traffic*** directly relates to the connectivity needs for the use cases. In chapter 3, we have seen several categories of traffic that have different QoS requirements:

- Streaming of 360 VR video content requires a low latency (of 20-40ms), which is substantially lower than required for many other applications (including many other types of streaming video).
- The public safety communication applications require a higher reliability of the connectivity service than other applications.

- Some of the automated driving applications combine very stringent requirements on latency (as low as 10-25 ms) with a need for a very high reliability of the connectivity service.

It is important to note that the condition in the Regulation revolves around specific categories of *traffic* rather than specific categories of *services* or *applications*. As seen earlier in the VR and autonomous driving cases, not all the traffic flows from an application necessarily have the same stringent requirements. In the VR case, for example, the traffic used in the purchase of the content and the initial downloading of the app does not require the low latency needed for the 360 VR streaming. A straightforward interpretation of the condition, in this case, would be that the QoS differentiation allowed for the 360 VR streaming traffic is not allowed for the traffic used in the purchase and download. In itself, this leads to a targeted application of QoS differentiation. It does lead to the practical question to what degree the traffic with less demanding requirements needs to be separated from flows for the specific categories of traffic. In the VR case, for example, the 360 VR content can be expected to involve a much larger amount of data than the initial download of the application. In practice, separating the traffic streams in different categories may therefore only have a small effect.

5.5.2 Transparency and non-discrimination

The assessment of the ***transparent*** and ***non-discriminatory*** use of QoS differentiation will strongly benefit from the use of standardised 5QI (5G QoS Identifier) values. The 5QI values bring transparency into the traffic management measures applied by the mobile operator. The non-discriminatory use can be assessed if it is known which 5QI values are used for which traffic flows, as it can then be analysed whether traffic flows with equivalent QoS requirements are indeed handled in the same way. The use of signalled QoS characteristics instead of standardised 5QI values would require a mobile operator to describe the combination of (more detailed) QoS parameters applied to specific traffic categories. In practice, this will involve more work for the mobile operator and also for NRAs and CAPs that want to have an insight into the operator's traffic management.

The condition that the traffic management ***shall not monitor the specific content*** is an important point. The Regulation (recital 9) and the Guidelines (paragraphs 64 and 70) explain that the mobile operator must base the assignment of traffic flows to traffic categories exclusively on information provided by the service and application. The operator is not allowed to monitor specific content (e.g., a 360 VR stream) carried as payload in the traffic to obtain information for the assignment, while the use of information available at the IP and TCP level is allowed. In general, this makes it challenging for a mobile operator to guarantee that all traffic flows with equivalent QoS requirements are indeed assigned to the same traffic category, in particular when an application involves multiple flows with different QoS requirements. The CAP can assist the operator by marking the traffic flows, for example through DiffServ Code Points (DSCPs) that can be mapped to 5QIs by the operator. However, the DSCP values may be changed in intermediate networks in the internet as there is no common practice to use them end-to-end. More fundamentally, it does not remove the dependency of the mobile operator on information from external sources for the correct assignment of traffic flows to

categories. An example scenario where this is a problem is when one CAP provides (standardised) 5QI values for its traffic flows, while another CAP does not provide specific information.

5.5.3 Duration of QoS differentiation

The condition that the QoS differentiation ***shall not be maintained for longer than necessary*** is at first sight at odds with the fact that the QoS requirements of a given traffic category are stable over time. A straightforward approach would, therefore, be to apply the QoS differentiation permanently. The *actual effect* of QoS differentiation may vary over time and in particular on the network load. The phrasing in the Regulation only allows the application of QoS differentiation when it is needed (and thus has an actual effect). The Guidelines do not rule out the permanent use of traffic management measures (such as QoS differentiation) but do explicitly raise the question whether this is reasonable (paragraph 73).

In practice, QoS differentiation involves (at least) two functions:

1. Marking of traffic, for example through 5QI values.
2. Acting on the markings, for example through the policies in routers and schedulers.

From an operational point of view, it may be convenient for a mobile operator to permanently mark the traffic, but only activate the QoS policies when needed. This would probably be allowed under the Guidelines as it can be interpreted as an implementation of the trigger function in paragraph 73.

5.5.4 Conclusion: medium to high complexity

Based on the analysis above, we expect that the analysis of QoS differentiation in practical cases has a medium to high complexity.

5.6 Topic 3: Local access to the internet

Local access to the internet as depicted on the left-hand side of Figure 29 represents a change to the topology of the internet. It can be interpreted as the internet getting bigger, in the sense that it reaches locations closer to the end users, leaving a smaller stretch to be bridged by the IAS. When it is combined with a more local provision of services and applications through local servers, it can also be viewed as the internet getting smaller, as the traffic of end users traverses a smaller segment of the internet.

As local access changes the *network topology* of the internet rather than the treatment of the traffic on the internet, it does not introduce new issues in the application of the Regulation or Guidelines compared to situations without local access. We, therefore, rate the complexity of this topic as low.

A secondary issue that is on the sideline of the Regulation and Guidelines is the issue of interconnection of potentially smaller and more local networks in the internet that play a role in local access. Smaller networks need to be interconnected to other, typically larger networks to become part of the internet and, vice versa,

other networks need to connect to the smaller networks if applications and services need to be provided through local access. In local access situations, the number of networks available for interconnection may be smaller than at larger hubs or internet exchanges, potentially giving the owners of these networks a larger technical or commercial influence. Paragraph 6 of the Guidelines explains that NRAs may investigate the interconnection policies and practices of ISPs if they limit end-user rights related to internet access. BEREC has recently performed a study in this area [71].

5.7 Topic 4: Public and private services and associated networks

The demarcation of public and private services is a topic that policymakers and NRAs have dealt with earlier. It can be expected to be a recurring topic in the context of 5G networks, as 5G architecture options such as slicing will make it easier to create private networks. For example, slices 2, 3 and 4 in Figure 29 show a VPN that uses separate mobile connectivity rather than an underlying IAS as commonly done today. In the application of the Regulation and Guidelines, both the *service* and the (underlying) *network* configuration are relevant. Article 2 of the Regulation defines the IAS as "*a **publicly available** electronic communications service that provides **access to the internet**, and thereby connectivity to virtually all end points of the internet …*"

The public safety and autonomous driving use cases in this study show the importance of this definition. A similar analysis has been performed for the case of operational rail communications ([72], [73]).

The basic public safety case from section 3.3 is offered to a predetermined group of end users, including the police, fire department and emergency medical services. It is therefore not a publicly offered service. Furthermore, the basic service does not offer internet access. Therefore, two independent factors place the basic service outside the scope of the rules for IAS. But for both factors, there are reasonable scenarios that change or remove them:

- The group of end users to which the service is offered may grow as additional stakeholders with critical communication needs are added. Examples are company fire departments in larger industrial estates, airport security services and port authorities.
- The basis public safety service is extended with a priority internet access to cater for the information needs of emergency services personnel in situations where the intensified use of networks by the general public causes congestion.

It is worth noting that if the public safety service is outside the scope for IAS, the rules for SpS come into play that apply for both publicly and non-publicly offered services.

In the automated driving case, the group of end users to which the services are offered is also relevant. Depending on the services and the business model, the services can be offered to car owners. This means that the service is offered to the general public and is, therefore, a public offering. Alternatively, the service could be offered to a smaller group, such as car manufacturers or roadside operators.

Depending on whether these groups are considered to be predetermined, this could lead to qualification as a non-public offering. As the services and business models in automated driving develop, the target groups can change. For the services that provide access to the internet, the change in target group can thus change the applicability of the rules for IAS. In section 5.11 we analyse the issue of services that offer access to only a limited number of end points in the internet, which can be relevant in the automated driving case as well. This touches the phrasing "*connectivity to virtually all end points of the internet"* in the definition of IAS.

Based on the analysis above, we expect that the analysis of public and private services in practical cases has a low to medium complexity.

5.8 Topic 5. Objective requirements for SpS

The room for the offering of SpSs by mobile operators and CAPs is defined in Article 3.5 of the Regulation:

> "*Providers of electronic communications to the public, including providers of internet access services, and providers of content, applications and services shall be free to offer* ***services other than internet access services*** *which are optimised for specific content, applications or services, or a combination thereof, where the* ***optimisation is necessary in order to meet requirements of the content, applications or services for a specific level of quality.***"

As explained in the Guidelines paragraphs 99, 108 and 111, the benchmark for assessing the necessity of the optimisation is the quality that can be achieved over the standard best-effort delivery in IAS. Based on the use cases and the points analysed in previous sections in this chapter, a number of observations can be made.

5.8.1 Multiple IASs

In this study, we assume that there can be multiple IASs with different traffic management settings in one mobile operator network (section 5.4). This means that for the assessment of the quality achievable over IAS, it may be necessary to select one of the available IASs as the benchmark, based on, for example, its QoS characteristics or the proportion of the operator's customers that uses it. Alternatively, one can construct the benchmark IAS by combining characteristics from multiple IASs, with the disadvantage that the quality cannot be evaluated in practice. With multiple IASs, there may be a choice between IAS with and without QoS differentiation (section 5.5). It is important to note that QoS differentiation can lead to higher but also to lower quality for the content, application or service concerned, depending on the alignment of its QoS requirements and the traffic categories considered in the QoS differentiation.

A related question is whether an increase in quality through *potential* QoS differentiation in the IAS is incorporated in the assessment. In situations where a network operator already applies QoS differentiation for certain traffic categories, it could be considered whether the addition of a category with specific QoS requirements could increase the quality achievable over IAS to the required level. A

drawback of such an approach is that the achievable quality cannot be evaluated in practice and the mobile operator may not be prepared to offer it.

5.8.2 Absolute measure of required quality

The assessment of the necessity of the optimisation must be based on a *specific level of quality*. This implies the use of an absolute measure of quality. This is different from the assessment of QoS differentiation in IAS that is based on (relative) differences in QoS requirements between traffic flows. For some services, the specific level of quality can indeed be expressed in an absolute measure. A well-known example is the traditional two-way voice communication service for which the overall quality can be expressed as a Mean Opinion Score (MOS) [74]. Such MOS scales have been developed for some other services (such as broadcast video [75]) as well, but for many other services, they are not available. In the use cases in this study, there is no absolute measure for the quality at the service or application level. Instead, the focus is on the QoS requirements at the network level: such 100 Mbit/s bandwidth and 20-40 ms latency in the VR use case, and 10-100 ms latency and 99.99 to 99.999% reliability of the connectivity for the specific cases in automated driving. As illustrated by the use cases, the QoS requirements may still need to be expressed in ranges rather than point values.

5.8.3 Characterisation of quality achievable over the internet

The QoS requirements must compare to the performance characteristics of the best-effort internet, which can vary over time, location and between users. The monitoring and characterisation the quality of the IAS has been the subject of several studies by BEREC (e.g. [76], [77], [78]). Network operators also have their own data on the performance of their IASs. In principle, a characterisation of the quality achievable over best-effort internet requires a quantitative assessment of the performance of IAS, based on measurements or insights in the network dimensioning in the operator network. As quantitative analysis is outside the scope of this study, we restrict ourselves to an important qualitative point.

The assessment is relativity straightforward when the QoS requirements of the SpS are much stricter (say, by an order of magnitude) than can be achieved over typical high-end or mid-range IAS services offered in the market. In this situation, there is no need for a detailed analysis or measurement of IAS performance. An example here is provided by the very high requirement for the reliability of connectivity service in the public safety use case. In all likelihood, this requirement cannot be met by a typical IAS. If the difference between the service requirements and the performance offered by IASs becomes smaller, the assessment becomes more difficult as the characterisation of IAS needs to be more precise. With a more detailed analysis of IAS performance, additional questions can present themselves: would it be necessary to take into account the differences in IAS performance between different regions within a country? Would it be necessary to consider differences between IASs offered by multiple operators? Depending on how fine-grained the performance of IAS is analysed, one could have outcomes in which a service qualifies as specialised in one region (or operator network) but not in other regions (or operator networks).

5.8.4 Multiple traffic flows in a service

As already observed in the analysis of QoS differentiation in section 5.5.1, not all of the traffic flows from a service or application necessarily have the same stringent requirements. Some of the traffic flows, such as the traffic used in the VR use case for the purchase of the content and the initial downloading of the app can be carried over the best-effort internet while preserving the desired quality for the overall service. This leads to the question of what level of detail the assessment analysis needs to zoom in on different traffic flows or components of a service. Would it be sufficient that a key or major part of the service depends on a quality that cannot be achieved over the internet? Or would it require a breakdown of the overall service into smaller and more specific services to be analysed?

5.8.5 Comparable content

The Guidelines contain an additional relevant condition for the assessment of SpSs in paragraph 110: "*If assurance of a specific level of quality is objectively necessary, this cannot be provided by simply granting general priority over comparable content."* This touches on two points in the VR use cases in this study:

- In the case of local processing of the video streams, the content is typically unique for the participating end users. It is not comparable to other content, in the sense that similar content cannot be used in the service for the users involved.
- In the case of local caching, the content is typically available in other locations, but it is made available through more functions and steps than "simply granting general priority over comparable content".

Therefore, this condition is not expected to raise issues.

5.8.6 Conclusion: potentially high complexity

Based on the analysis above, we expect that in practical cases, the analysis of the objective need for optimisation in SpS can have high complexity.

5.9 Topic 6. Impact of SpS on IASs

In Article 3(5) of the Regulation, it is made very clear that the offering of SpS should not be at the cost of the quality of IAS: "*Providers … may offer or facilitate such services only if the network capacity is sufficient to provide them in addition to any internet access services provided. Such services … shall not be to the detriment of the availability or general quality of internet access services for end-users.*"

A somewhat implicit assumption in this article is that capacity and resources in a network are scarce. In 5G mobile networks, this can be expected to be the case in many situations as well, despite the substantial increase in capacity achieved through the new technologies applied in radio and core networks. Note that quality is not only dependent on bandwidth but also on other QoS parameters such as latency and jitter. For mobile networks, a specific point is made in the Regulation (Recital 17) and in the Guidelines (paragraph 123): a temporal negative impact of SpSs on the quality of IASs is acceptable, as the number of users in a (radio) cell

may be difficult to anticipate. The impact should be unavoidable, minimal and of short duration though.

In general, the condition made in Article 3(5) of the Regulation calls for a quantitative assessment of the effect of an SpS on the performance of (all) IASs in a network. Since the description of the use cases in this study does not include quantitative network planning and dimensioning, the analysis below is restricted to a number of generic, qualitative points. These points can be expected to be addressed in a quantitative assessment of the impact of a specific SpS.

As seen in Figure 28, the IASs that the Regulation aims to protect from degradation can be in the same slice (slice 2) or in other slices (such as slice 1) than the slice in which the SpS is provided. The distribution of capacity over different slices in a 5G network is determined by the slice capacity management functions that are currently under development in 3GPP (section 2.2.1). Furthermore, there can be multiple IASs with different traffic management settings in one slice, potentially in combination with SpSs. Here, the traffic flows can be prioritised and separated using QoS differentiation (section 2.2.3). Thus, at least two types of multiplicities can be encountered in the assessment, which each bring their own mechanisms for distribution of capacity. Figure 30 illustrates where the capacity management touches the services offered in the slices 1 and 2 from the (larger) consolidated architecture if SpS (c) is introduced. Note that the services in these slices are in general provided to many end users that share the capacity in the slices.

- The first and most obvious point to assess is whether the capacity of slice 2 is extended (orange arrow A in the figure). If the capacity of the slice is extended with the (predicted) capacity of SpS (c), then the capacity of slice 2 suffices. Note that the capacity of SpS (c) would be expected to grow with the take-up of the service after introduction, thus requiring a further extension of the capacity.
- Still, one would need to assess how the capacity in slice 2 is distributed between IAS (b) and SpS (c), indicated by orange arrow B. This could, for example, depend on the details of the traffic management applied in slice 2 through QoS differentiation (section 2.2.3).
- In addition, the effect of an extension of the capacity of slice 2 on the capacity of slice 1 needs to be assessed (arrow C). Here, it is important to note that the allocation of capacity to slices can be dynamic with minimum and maximum values per slice.

In a further quantitative step in the analysis, the absolute capacities for the IASs and the SpS and their relative proportions would be included to assess the impact of the SpS.

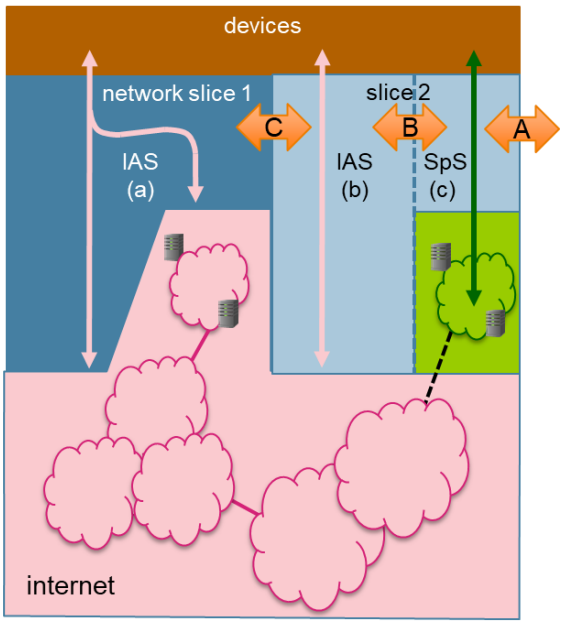

Figure 30. Effect of capacity management on the services offered in two slices.

In specific cases, the impact of the SpS on IASs may be of limited duration. An example is a temporal increase of the capacity for critical Public Safety communications during emergency situations. If there would be no corresponding temporal expansion of the total network capacity, this increase would occur at the expense of other services, including IASs, for other users.

The Regulation (recital 17) and Guidelines (paragraph 174-176) describe how NRAs can use measurements to monitor the quality of IASs and check whether the IAS indeed does not suffer from the offering of SpS. The measurements involve network-level parameters such as bandwidth, latency and jitter. BEREC has investigated such measurements in several studies ([76],[77],[78]). The Guidelines suggest the option to measure the IAS quality with and without the presence of the SpS under investigation (paragraph 121). This may be difficult to achieve in practice as it would require switching off the SpS, which is a service with a need for a specific quality. Measuring the quality of IAS before and after the introduction of the SpS does not have this difficulty. Here, the challenge is to isolate the effect of the SpS from the other factors affecting IAS performance, such changes in its number of end users and the take-up of new applications offered by CAPs.

In parallel to the measurements that NRAs can perform to assess the quality of IASs, mobile network operators will monitor the performance of their IASs and SPSs. They can be expected to use this as input for their network dimensioning and upgrades and to check the capacity estimates they have made before the introduction of an SpS. According to article 5(2) of the Regulation, NRAs can request information on traffic management and network capacity from operators. The level of detail in the information to be provided can be determined by the NRA. In principle, an NRA can thus have the same inside perspective on network dimensioning, capacity allocation and traffic management as the mobile network operator. They can combine this view with their own independent measurements to assess whether the condition for offering the SpS is met. In practice, this may still be a complex exercise, given that network and capacity management is a major

and non-trivial activity for mobile operators, with variations over time and geography.

Based on the analysis above, we rate the relative complexity of the analysis of this topic as high.

5.10 Topic 7. SpS and connections to the internet

In several architecture options for the use cases in this study, there is a need to connect IP Data Networks to the internet. These connections are indicated by the dashed lines in the consolidated architecture (Figure 27). Examples are:

- VR use case: the connectivity used to download content into local caches from the internet (in the Slicing & Service hosting architecture option);
- PS use case: the connectivity used to provide internet access from the public safety slice (in the Slicing architecture option);
- AD use case: the connectivity used to exchange information between the local ITS cloud and servers in the internet (in the Local access & Slicing option)

If the services and applications supported in these architecture options are to comply with the rules for SpSs, as indicated in Figure 28, then this connectivity seems to be at odds with the Regulation and Guidelines. First, the Regulation defines SpSs as *other than IAS* in Article 3(5). Second, the Guidelines state in paragraph 110 that *specialised services do not provide connectivity to the internet.*

Paragraph 115 of the Guidelines describes a more subtle interpretation though: "*VPNs could qualify as specialised services in accordance with … the Regulation. However, … to the extent that corporate services such as VPNs also provide access to the internet, the provision of such access to the internet by a provider of electronic communications to the public should comply with Article 3(1) to (4) of the Regulation."* In this interpretation, an SpS is allowed to provide connectivity to the internet, as long as this occurs through a separate IAS that complies with all rules for IAS (Figure 31, left). This sequence of an SpS and an IAS is found in the public safety use case example.

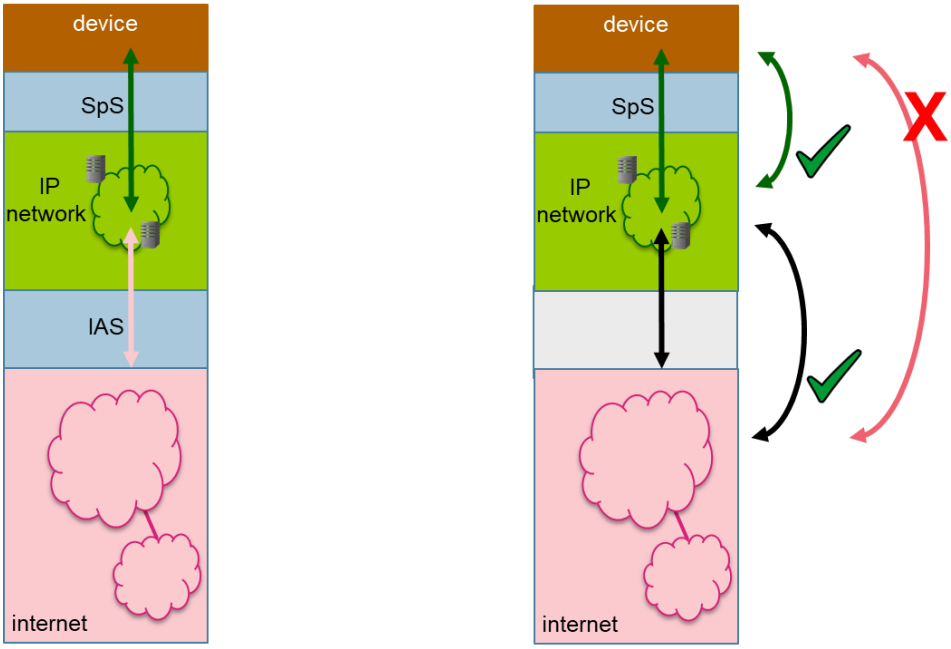

Figure 31. Access to the internet from an SpS through a separate IAS (left) and connectivity to the internet used to connect servers in the IP network to servers on the internet (right).

The situations in the VR and the automated driving use cases are different. Here, the purpose of the connectivity to the internet is to connect one or more servers in the VR and AD slices to servers on the internet. The purpose is not to provide internet access to the end user device. The device can connect to the IP data network through the SpS (the green leg in right-hand side of Figure 31), the servers in the IP data network can connect to servers on the internet (the black leg), but the two legs are not connected and therefore do not provide internet access to the device. This situation complies with Guidelines paragraphs 102, 126 and 127 that describe that specialised services are not usable or offered as a replacement for IAS.

We expect the complexity of the analysis of this topic to be low. The analysis is relatively straightforward once the details of the architecture options, the Regulation and the Guidelines are carefully combined.

5.11 Topic 8. Connectivity to a limited number of internet end points

The autonomous driving use case provides a situation where it may be useful to have a service that provides connectivity to only a limited number of end points on the internet. For example, the motor management system in a car may be configured to only connect to specific servers of the car manufacturer or component supplier on the internet. The access to a limited number of internet end points is addressed in the Guidelines (see section 2.4.3), which describes a potential exception to the overall ban on sub-internet services. The exception is made for situations where the number of reachable end points is limited by the nature of the device, rather than by the mobile operator that provides the connectivity service. This is expected to be relevant for Internet of Things (IoT) devices with relatively limited functions and connectivity needs.

In the Guidelines, the description of the exception involves elements from the preceding sections in this chapter: IAS, SpS, private networks and devices, illustrating the complexity of this point. The analysis in this section focusses on the example from the automated driving use case. If the motor management system is configured to connect only to specific servers (and therefore specific end points) in the internet, the connectivity needed for this can be still provided by a regular IAS. Alternatively, it can be provided by a combination of an SpS and an IAS, as seen earlier in Figure 31 (left). The description in paragraph 18 becomes relevant when the set of reachable end points is limited in the connectivity service as well, in addition to the limitations already introduced by the device. The car manufacturer and the mobile operator may prefer this approach for traffic management, security or other reasons. The resulting connectivity service is a sub-internet service, but one that does not introduce limitations in the internet access that have not already been introduced by the device. An NRA could thus argue that this sub-internet service is not used to circumvent the Regulation. Alternatively, the NRA could argue that sub-internet services are clearly not allowed by the Regulation. The traffic management, security or other requirements would then need to be addressed through other mechanisms.

We expect that the analysis of this topic in practical cases will be of medium complexity, compared to the other topics in this report.

5.12 Topic 9. Access control

In the public safety use case, active access control can be used to ensure the availability of sufficient capacity during network congestion, typically associated with high network loads during (large-scale) emergency situations (section 4.3.1). Ensuring or even temporarily expanding the capacity for critical Public Safety communications will likely occur at the cost of the capacity for other services, including IASs in the network. As discussed in section 5.9, if the critical Public Safety services are considered to be offered as an SpS, then this represents an impact of an SpS on IASs.

Aside from this, the use of access control needs to comply with step 3 of the rules on traffic management for IAS:

> "*Providers of internet access services .. shall not block … interfere with, degrade or discriminate between specific content, applications or services, except as necessary … in order to… prevent impending network congestion and mitigate the effects of exceptional or temporary network congestion, provided that equivalent categories of traffic are treated equally.*"

Looking further at how active access control affects the IASs in the network, it is seen that the rules do not introduce issues:

- The barring (which corresponds to blocking) is only applied during temporary congestion situations, associated with unpredictable emergency situations.
- The barring occurs at a coarse level where a distinction is made between service categories such as SMS, emergency calls and IAS. The barring does not involve any differentiation between traffic flows within the IAS. Therefore, equivalent categories of traffic within IAS are intrinsically treated equally.

We expect the analysis of this topic to be of low complexity.

6 Conclusion and recommendations

6.1 Conclusions

The technological neutrality of the Regulation allows 5G network technology itself to develop. There is no a priori ban on any 5G technology ingredient.

Our analysis underlines the importance of technological neutrality. This is a well-established principle that is adhered to in the Regulation and the Guidelines. It plays a crucial role in the analysis. What matters for the compliance with net neutrality rules is how the 5G technologies are used to support services and applications, rather than the technologies themselves. Therefore, the European net neutrality rules do not introduce a ban on any 5G technology ingredient, also not on the technologies that are being developed with the aim to differentiate between traffic flows and applications.

The assessment of the alignment of 5G with net neutrality rules depends not only on the 5G technologies, but also on the specific combination of services, applications and network architecture. It is not possible to come to an overall assessment with a single outcome on the alignment of 5G technology with net neutrality rules.

The central question in the assessment of the compliance with net neutrality rules is whether the services and applications supported by the 5G technology components adhere to the conditions and rules for IASs and SpSs, whichever are applicable. It is these conditions and rules that determine the room for mobile operators and CAPs in their use of 5G technology.

In our analysis, slicing provides a relevant illustration of this point. Slicing is a key 5G technology for mobile operators to support tailored connectivity to different services and applications. The use of slicing will vary, as illustrated by the figures in the previous chapter. In 5G architectures that use slicing, an IAS is always in a slice. A slice can be used exclusively to provide an IAS. Alternatively, a single slice can be used to simultaneously provide an IAS and an SpS. A slice can also be exclusively used to provide an SpS. Thus, the use of slicing technology in a mobile operator network can bring in the rules for IAS, SpS or both, depending on the services and applications that are supported. It is not possible to come to an overall assessment with a single outcome on the alignment of slicing with net neutrality rules. The topics that are encountered in the assessment and the outcome depend not only on the 5G technology, but also on the specific combination of services, applications and network architecture. This is true for slicing, but also for other key 5G technologies such as QoS differentiation. A consequence is that mobile operators, content and application providers and national regulatory authorities will need to do further analysis to evaluate whether a particular type of (tailored) connectivity complies with the net neutrality rules.

The topics encountered in the assessment of the compliance are of varying complexity. The impact of Specialised Services on Internet Access Services *and* the objective need for optimisation in Specialised Services *are expected to have the highest complexity.*

Based on our analysis of the three use cases (VR in Media, Autonomous Driving and Public Safety) and the key 5G technology ingredients, we have identified nine topics that are relevant to the assessment. We have positioned these topics in a consolidated 5G architecture that shows typical situations in which they come into play. The topics are summarised in Table 7, together with our expectation for their relative complexity, based on the analysis in the previous chapter.

Table 7.Summary of topics in the alignment of 5G architecture options with net neutrality rules and their expected relative complexity.

Topic	**Key points identified in the analysis**	**Relative regulatory complexity**[11]
Multiple IASs with different traffic management settings	• Interpretation of *sender and receiver* in Art 3.3 of the Regulation • Note: assumption needed in remainder of analysis - it is allowed to have multiple IASs with different traffic management settings for a given end user	low
QoS differentiation within IAS	• Applications with multiple different traffic flows • Transparency through 5QI values or other methods • Dependency of ISP on other entities for assignment of traffic flows to traffic categories • Duration of QoS differentiation	medium to high
Local access to the internet	• (potentially:) IP interconnection of local networks	low
Public and private services and associated networks	• Size and scope of predetermined group of end users in private service	low to medium
Objective need for optimisation in SpS	• Determination of IAS for benchmark in case of multiple IAS offers • Variation of IAS performance between geographical regions and operators • Services comprising multiple traffic flows	high, except if SpS requirements are clearly much stricter than achievable over IAS.
Impact of SpS on IASs	• Multiple IASs affected by one SpS, within and outside the slice used for the SpS.	high

[11] We define this as the relative complexity expected to be encountered by national regulatory authorities, mobile operators, and content and application providers when they analyse specific cases with more context information and (quantitative) details than the use case-inspired analysis made here (section 5.3).

	• Isolation of the effect of the SpS on IAS from other effects occurring in mobile network at the same time • Complexity of network and capacity management in mobile network with many services and applications in general	
SpS and connections to the internet	• Connectivity to internet from SpS through separate IAS • Connectivity between different legs between end user device and internet	low
Connectivity to limited number of internet end points	• Evaluation whether sub-internet service is acceptable for providing connectivity in specific situations	medium
Access control	(no issues if use is restricted to network congestion in emergency situations)	low

In our analysis, we found that several topics that appear to be complex at first sight, such as *SpS and connections to the internet*, become relatively straightforward to assess once the details of the architecture options, the Regulation and the Guidelines are carefully combined. We expect that the low to medium complexity topics lend themselves to the formulation of “rules of thumb” within NRAs, network operators or CAPs. They can be formulated based on internal analysis or, at a later stage, be derived from earlier assessments by NRAs and the resulting case law.

Other topics such as the *impact of SpS on IASs* can be expected to remain relatively complex. There are no fundamental problems that prohibit their analysis. However, the complexity of these topics is likely to make them unsuitable for a generic “rule of thumb” approaches. They require a case-by-case approach. The complexity depends on the level of detail that NRAs, network operators and CAPs pursue in their analyses.

The topics encountered in the assessment are relevant for services and applications provided over mobile and fixed networks in general. They are not exclusively related to 5G technology.

A final observation is that the topics identified as relevant in the assessment are not exclusively related to 5G. They can also present themselves in the analysis of services and applications provided over 3G, 4G and pre-5G networks. As the Regulation and Guidelines are to a (very) large extent technology neutral, the analysis of the topics would be largely similar. The topics can be expected to be more relevant in 5G networks though, as 5G technology provides more extensive support and flexibility for tailored mobile connectivity aimed at specific sectors or user groups.

6.2 Recommendations

Our first recommendation is to clearly distinguish between 5G architecture elements on the one hand and the net neutrality concepts of IAS and SpS on the other. One should keep a technology-neutral view and not attempt to define a one-to-one mapping between the two. Two important examples of this are:

1. A slice is not the same as an SpS. As already explained above, slicing can be used to support an IAS, an SpS or both.
2. The application of QoS differentiation is not limited to IAS. QoS differentiation can be used as a method for traffic management within an IAS. However, it can also be used to assure the quality of SpSs. A prominent example of the latter is the VoLTE architecture in 4G networks.

Our second recommendation is that subject matter experts at national regulatory authorities, mobile operators, and content and application providers build upon our approach and findings in their assessments. We expect that the consolidated architecture model provides a good starting point to structure the overall discussion on services and applications over 5G networks and their compliance with net neutrality rules. For the analysis of specific services and applications, the three-step approach applied to the use cases in this report is recommended:

1. Determine the connectivity requirements of the services and applications in the use case.
2. Develop the 5G architecture options to support the connectivity requirements. The 5G technology ingredients described in this report are expected to play an important role here.
3. Evaluate the alignment of the combination of services, applications and architecture options with net neutrality rules. Here, the analysis of the specific topics made in this report can probably (partly) reused.

Mobile operators, content and application providers and national regulatory authorities can use this approach to develop their own individual analysis. These steps can also be used to structure the discussion among stakeholders and come to a shared analysis. Such a shared analysis would be beneficial for providing clarity and reducing uncertainties that industry may encounter in its development of roadmaps for 5G networks and applications that rely on tailored connectivity.

References

[1] ITU-R Recommendation M.2083-0, Framework and overall objectives of the future development of IMT for 2020 and beyond, September 2015

[2] Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access (2015)

[3] BEREC Guidelines on the Implementation by National Regulators of European Net Neutrality Rules. BEREC report BoR (16) 127, August 2016

[4] Evolution of Cellular Technologies, Chapter 1 of, Fundamentals of LTE, Arunabha Ghosh, et al., Prentice Hall, August 2010, ISBN-13: 978-0-13-703311-9

[5] The Evolution of Mobile Networks: 1G 2G 3G 4G and 5G, http://www.kosbit.net/evolution-mobile-networks-1g-2g-3g-4g-5g/

[6] Wegwijs in frequentieland, Geschiedenis, Autotelefonie, http://www.frequentieland.nl/geschiedenis/autotelefoon.htm

[7] ITU-R Recommendation M.687-2, International Mobile Telecommunications-2000 (IMT-2000), February 1997

[8] ITU-R Recommendation M.816-1, Framework for services supported on International Mobile Telecommunications-2000 (IMT-2000), March 1992

[9] ITU-R Recommendation M.1457-1, Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-2000 (IMT-2000), August 2001

[10] ITU-R Recommendation M.1645, Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000, March 2003

[11] ITU-R Recommendation M.2012, Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications Advanced (IMT-Advanced), January 2012

[12] 3GPP TS 23.501, System Architecture for the 5G System, Version 15.0.0, December 2017

[13] 3GPP TS 22.261, Service requirements for next generation new services and markets, Version 16.2.0, January 2018

[14] 3GPP TS 28.530, Management of network slicing in mobile networks; Concepts, use cases and requirements, Version 0.5.0, February 2018

[15] 5G Network Transformation, 5G Americas White Paper, December 2017, http://www.5gamericas.org/files/3815/1310/3919/5G_Network_Transformation_Final.pdf

[16] 3GPP TR 28.801, Study on management and orchestration of network slicing for next generation network, Version 15.1.0, January 2018

[17] 3GPP TS 38.300, NR; Overall description; Stage-2, Version 15.0.0, January 2018

[18] Understanding Information Centric Networking and Mobile Edge Computing, 5G Americas, December 2016

[19] Official Document IR.92 - IMS Profile for Voice and SMS, GSM Association, Version 11.0, 15 June 2017

[20] Official Document IR.94 - IMS Profile for Conversational Video Service, GSM Association, Version 12.0, 12 June 2017

[21] Marsden, Chris. Network neutrality from Policy to Law to Regulation, Manchester University Press. 2017.

[22] A view of traffic management and other practices resulting in restrictions to the open Internet in Europe - Findings from BEREC's and the European Commission's joint investigation, BoR (12) 30, 29 May 2012.
[23] Pieter Nooren, Andra Leurdijk, Nico van Eijk (2012),"Net neutrality and the value chain for video", info, Vol. 14 Iss: 6 pp. 45 – 58, http://dx.doi.org/10.1108/14636691211271226
[24] Pieter Nooren et al. Regulation in the converged media-internet-telecom value web. October 2014. Report. Retrieved November 17, 2017 from http://publications.tno.nl/publication/34611843/NhocfJ/TNO-2014-R11482.pdf
[25] Tim Wu, (2003), ''Network neutrality, broadband discrimination'', Journal of Telecommunications and High Technology Law, Vol. 2, pp. 141-79.
[26] Preserving Internet Freedom: Guiding Principles for the Industry, Michael K. Powell, February 8, 2004, http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-243556A1.pdf
[27] FCC, Report and Order, In the Matter of Preserving the Open Internet; Broadband Industry Practices; GN Docket No. 09-191, WC Docket No. 07-52, December 23, 2010, Federal Register, Vol. 76, No. 185, p. 59192, 23 September 2011
[28] FCC. Protecting and Promoting the Open Internet, GN Docket No. 14-28, Report & Order on Remand, Declaratory Ruling, and Order, 30 FCC Rcd. 5601 (2015)
[29] Statement of Chairman Ajit Pai on Restoring Internet Freedom, FCC, 4 January 2018, WC Docket No. 17-108, https://apps.fcc.gov/edocs_public/attachmatch/FCC-17-166A2.pdf
[30] To kill net neutrality, FCC might have to fight more than half of US states, February 16, 2018, https://arstechnica.com/tech-policy/2018/02/to-kill-net-neutrality-fcc-might-have-to-fight-more-than-half-of-us-states/
[31] Directive 2009/140/EC (Framework Directive) of the European Parliament and of the Council, November 25, 2009
[32] Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services (Universal Service Directive) OJ L 108/51 (24 April 2002).
[33] Wijziging van de Telecommunicatiewet ter implementatie van de herziene telecommunicatierichtlijnen, Amendement van het lid Verhoeven c.s., Tweede Kamer der Staten-Generaal, vergaderjaar 2010-2011, 32 549, Nr 29 (in Dutch).
[34] Sky VR gives fans virtual reality journey into sporting events, October 5, 2016, http://www.skysports.com/football/news/12040/10604000/sky-vr-gives-fans-virtual-reality-journey-into-sporting-events
[35] BT Sport to broadcast UEFA Champions League final in 360° VR, May 22, 2017, http://home.bt.com/tech-gadgets/phones-tablets/bt-sport-to-give-away-virtual-reality-headsets-for-free-ahead-of-uefa-champions-league-final-11364181144704
[36] Hong Kong Sevens Tournament Broadcasted Live In 360-Degree Virtual Reality, May 2, 2017, https://www.sporttechie.com/hong-kong-sevens-tournament-broadcasted-live-in-360-3d/
[37] NextVR - Get exclusive access to sports, music, and entertainment immersive experiences in virtual reality, NextVR, https://www.nextvr.com/, retrieved February 2, 2018

[38] Facebook Spaces, https://www.facebook.com/spaces, retrieved February 2, 2018
[39] Facebook finally makes a virtual reality world, April 18, 2017, http://money.cnn.com/2017/04/18/technology/facebook-f8/index.html
[40] Eerste indruk: VR in Facebook Spaces geen eenzame ervaring meer, (in Dutch), April 19, 2017, http://www.nu.nl/reviews/4629759/eerste-indruk-vr-in-facebook-spaces-geen-eenzame-ervaring-meer.html
[41] Amazon is creating a new virtual reality platform, March 8, 2016, https://www.theverge.com/2016/3/8/11177854/amazon-vr-platform-job-posting
[42] Coming Soon to Amazon: Shopping in Virtual Reality? February 1, 2018, https://www.inc.com/kevin-j-ryan/amazon-wants-you-to-shop-in-virtual-reality.html
[43] The Zettabyte Era: Trends and Analysis, Cisco White Paper, 2017. Retrieved November 17, 2017, from http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.pdf
[44] Augmented reality: a class of displays on the reality-virtuality continuum, Paul Milgram et al, Proc. SPIE 2351, Telemanipulator and Telepresence Technologies, (21 December 1995); doi: 10.1117/12.197321; http://dx.doi.org/10.1117/12.197321
[45] Microsoft HoloLens, https://www.microsoft.com/en-us/hololens, retrieved 2 February 2018
[46] Xavier P. Burgos-Artizzu et al, Real-time expression-sensitive HMD face reconstruction, SIGGRAPH Asia, Kobe, Japan, November 2 - 6, 2015, Technical Briefs Article No. 9
[47] Netflix - Internet Connection Speed Recommendations, https://help.netflix.com/en/node/306, retrieved 2 February 2018
[48] YouTube Help - Live encoder settings, bitrates and resolutions, https://support.google.com/youtube/answer/2853702?hl=en-GB, retrieved 2 February 2018
[49] D. Podborski et al, Virtual Reality and Dash, International Broadcasting Convention, Amsterdam, 13-17 September 2017, https://www.ibc.org/download?ac=3809
[50] R. van Brandenburg, R. Koenen, D. Sztykman, CDN Optimization for VR Streaming, International Broadcasting Convention, Amsterdam, 13-17 September 2017.
[51] Draft Guidelines, VR Industry Forum, December 22, 2017, http://www.vr-if.org/wp-content/uploads/VRIF-Integrated-Guidelines-vrif2017.106.18-CES.rk_.pdf
[52] Virtual Reality (VR) media services over 3GPP, 3GPP TR 26.918 V15.1.0 (2017-12), http://www.3gpp.org/ftp//Specs/archive/26_series/26.918/26918-f10.zip
[53] VR and AR pushing connectivity limits, White paper, Qualcomm Technologies, May 2017, https://www.qualcomm.com/media/documents/files/vr-and-ar-pushing-connectivity-limits.pdf
[54] Latency – the sine qua non of AR and VR, Michael Abrash, Blogpost, December 29, 2012, http://blogs.valvesoftware.com/abrash/latency-the-sine-qua-non-of-ar-and-vr/

[55] What 5G could mean for broadcast, IBC365, 16 August 2017, https://www.ibc.org/tech-advances/what-5g-could-mean-for-broadcast/2141.article
[56] Website ETSI with information about TETRA etsi.org/technologies-clusters/technologies/tetra
[57] Official website of FirstNet USA firstnet.gov
[58] Website of PCP project Broadmap broadmap.eu
[59] Facilitering missie-kritisch mobiel breedband voor het OOV-domein, TNO Report R11193, October 2017.
[60] Research into linkages between the 700 MHz, 1452-1492 and 2100 MHz bands. Study conducted by Aetha, October 2016.
[61] consilium.europa.eu/en/council-eu/preparatory-bodies/law-enforcement-working-party/
[62] Website ITS programma Rijksoverheid (Ministerie I&M) beterbenutten.nl
[63] Website Europees ITS harmonisatieproject Intercor intercor-project.eu/
[64] Website Europees ITS standaardisatieplatform C-ROADS c-roads.eu/platform.html
[65] Informatie over EATA erticonetwork.com/european-automotive-telecom-alliance-presents-automated-driving-roadmap/
[66] Website Finse 5GSAFE project 5gsafe.fmi.fi/
[67] Website Europees (CEF) project CONCORDA ec.europa.eu/inea/en/connecting-europe-facility/cef-transport/concorda
[68] Website Europese IoT Large Scale Pilots (H2020) european-iot-pilots.eu/
[69] G. Pocovi, Automation for on-road vehicles: Use Cases and requirements for radio design, Aalborg University,
[70] Quick scan of selected mobile services and their qualification as specialised service in the context of the Dutch net neutrality rules, Pieter Nooren, Toon Norp, Annemieke Kips, Bram van den Ende, TNO report 2015 R10194, 23 March 2015
[71] BEREC Report on IP-Interconnection practices in the Context of Net Neutrality, BEREC draft BoR (17) 111, 1 June 2017
[72] Implications of Bearer Independent Communication Concept, European Union Agency for Railways, Report ERA 2016 17 RS, 25 July 2017
[73] J. Scott Marcus & Gabor Molnar, Network Sharing and 5G in Europe: The Potential Benefits of Using SDN or NFV, Digiworld Economic Journal, No. 108, 4th quarter 2017
[74] Methods for subjective determination of transmission quality, Recommendation ITU-T P.800 (08/96)
[75] Methodology for the subjective assessment of the quality of television pictures, Recommendation ITU-R BT.500-13 (01/2012)
[76] Net Neutrality Regulatory Assessment Methodology, BEREC draft BoR (17) 112, 1 June 2017
[77] Monitoring quality of Internet access services in the context of net neutrality, BEREC report BoR (14) 117, 25 September 2014

[78] Feasibility study of quality monitoring in the context of net neutrality, BEREC report BoR (15) 207, 30 November 2015

A Acknowledgements

The authors of this report thank the external experts that have contributed their views on 5G and Net Neutrality in interviews. The analysis in the report has greatly benefitted from their inputs on use cases and specific points in net neutrality. As can be expected, the focus and sometimes also the interpretation of the issues varies between experts. We, therefore, emphasise that the analysis presented in this report is that of the authors and not necessarily that of the experts.

Igor Curcio	Principal Scientist, Video Coding & Transport Nokia Technologies
Rob Frieden	Professor of Telecommunications and Law, Penn State University
Luca Belli	Senior Researcher, Center for Technology & Society, FGV Rio de Janeiro and Chercheur Associé, Centre de Droit Public Comparé, Université Paris 2
Tiia Ojanpera	Project Manager 5G Safe, VTT, Finland
Tomas Jacimavicius Michele Zarri Mani Manimohan	Director of Policy, Government & Regulatory Affairs GSMA; Technical Director, GSMA; Senior Director of Public Policy, GSMA
Scott Marcus	Senior Fellow, Bruegel
Stefan Christiernin David Lindgren Urban Friberg	Research Affairs and Innovation manager, NEVS; Car Connectivity Manager, NEVS; Lead Engineer Telematic, NEVS
Rob Koenen	President, VR Industry Forum

We also thank the contact persons at our project sponsors for their support in setting up the project. They were instrumental in creating the right multi-stakeholder environment for our analysis.

Mischa Prinsen Lubna Safeer Brenda van der Wal	Ministry of Economic Affairs and Climate Policy
Michiel van Dijk Diewertje Heermans Dennis Brouwer Onno Mantel	Authority for Consumers and Markets (ACM)
Paul Knol Robert van Erp	KPN
Joepke van der Linden Richard Marijs Han van Bussel	T-Mobile Netherlands
Patrick Blankers Simon Biemond	Ericsson
Anne van Otterlo	Nokia
Jurjen Veldhuizen	Huawei Technologies
Liesbeth Holterman	FME

B Abbreviations

3GPP	Third Generation Partnership Project
3/4/5G	Third, Fourth and Fifth Generation
5QI	5G QoS Identifier
AC	Access Control
AF	Application Function
AI	Access Identity
AMF	Access and Mobility Function
ARP	Allocation and Retention Priority
AS	Application Server
AUSF	Authentication Server Function
ATF	Auto TeleFoon
CAGR	Cumulative Aggregate Growth Rate
CAP	Content and Application Provider
CCBG	Critical Communications Broadband Group
DSCP	DiffServ Code Points
EATA	European Alliance Telecommunications and Automotive
eMBB	Enhanced Mobile BroadBand
EU	European Union
BEREC	Body of European Regulators for Electronic Communications
DN	Data Network
EPS	Evolved Packet System
ETSI	European Telecommunication Standards Institute
FCC	Federal Communications Commission
GSMA	GSM Association
HMD	Head Mounted Device
HSDPA	High Speed Data Packet Access
IAS	Internet Access Service
IMS	IP Multimedia System
IMSI	International Mobile Subscriber Identity
IoT	Internet of Things
IP	Internet Protocol
ISP	Internet Service Provider
ITS	Intelligent Transport System
ITU	International Telecommunication Union
LTE	Long Term Evolution
MEC	Mobile Edge Computing
mIOT	Massive IoT
MOS	Mean Opinion Score
MVNO	Mobile Virtual Network Operator
NEF	Network Exposure Function
NN	Net Neutrality
NGMN	Next Generation Mobile Networks
NR	New Radio (5G)
NRA	National Regulatory Authority
NRF	Network function Repository Function
NSSF	Networks Slice Selection Function

OEM	Original Equipment Supplier
OTT	Over The Top
PAMR	Public Access Mobile
PCF	Policy Control Function
PCP	Pre Competitive Procurement
PCRF	Policy and Charging Rules Function
PDU	Packet Data Unit
PPDR	Public Protection and Disaster Relief
PPI	Public Procurement of Innovation Solutions
PTT	Push-To-Talk
QoS	Quality of Service
QCI	QoS Class Identifier
QFI	QoS Flow ID
RAN	Radio Access Network
SD	Slice Differentiator
SIM	Subscriber Identity Module
SST	Slice Service Type
SMF	Session Management Function
SpS	Specialized Service
SUPI	Subscriber Permanent Identifier
TCCA	TETRA Critical Communications Association
TETRA	TErrestrial TRunked RAdio
TS	Technical Specification
UDM	Unified Data Management
UE	User Entity
uRLLC	Ultra-Reliable Low Latency Communication
UPF	User Plane Function
USIM	Universal Subscriber Identity Module
V2I	Vehicle to Infrastructure
V2N	Vehicle to Network
V2V	Vehicle to Vehicle
ViLTE	Video over LTE
VoLTE	Voice over LTE
VPN	Virtual Private Network

C Further elaboration of 5G QoS

This annex provides a further elaboration of the QoS mechanism in 5G and, where applicable, a comparison to the 4G mechanism.

- The smallest granularity for QoS is the **QoS Flow**. The corresponding 4G term is an IP-CAN Bearer;
- A **PDU Session** contains one or multiple QoS Flows. A PDU Session corresponds to an IP-CAN Session in 4G;
- A QoS Flow is identified by a QoS Flow ID (**QFI**);
- A QFI refers to a **QoS Profile**.
- A QoS Profile contains **QoS Parameters** comprising:
    - 5G QoS Identifier (**5QI**). This corresponds to the 4G QoS Class Identifier – QCI. For 5G there is a standardized 5QI table (see below) mapping the standardized 5QI values to corresponding QoS Characteristics (see below).
    - Allocation and Retention Priority (**ARP**). This corresponds to the same concept in 4G and determines whether or not a certain 'bearer' – QoS Flow can pre-empt other 'bearers', i.e. push other 'bearer' off the networks in case of resource needs;
    - Flow Bit Rate (**GFBR** – Guaranteed Flow Bit Rate, **MFBR** – Maximum Flow Bit Rate) applies only to Guaranteed Bit Rate QoS Flows, and specify the bit rate requirements of the GBR QoS Flow. They correspond to Guaranteed Bit Rate/ Maximum Bit Rate, respectively in 4G networks;
    - Aggregate Bit Rate (**Session-AMBR** – Session Aggregate Maximum Bit Rate, **UE-AMBR**) applies only to non-Guaranteed Bit Rate QoS Flows and specify the overall total bit rate requirements of all non-GBR QoS Flows combined. They correspond to APN-AMBR/ UE-AMBR, respectively in 4G networks;
    - Maximum Packet Loss Rate applies only to GBR QoS Flows and specifies the maximum packet loss rate that can be tolerated for the QoS Flow. There is no 4G equivalent for this;
    - Reflective QoS Attribute (**RQA**) applies only to non-GBR QoS Flows and specifies the availability of Reflective QoS. Reflective QoS enables the UE to map uplink traffic to a QoS Flow based on the QoS Flow used for the downlink. There is no 4G equivalent for this;
    - Notification Control applies only to GBR QoS Flows and specifies whether notifications are requested from the RAN when GFBR can no longer be fulfilled for a QoS Flow. There is no 4G equivalent for this;
- **QoS Characteristics** are associated to a 5QI; they are "guidelines for setting node specific parameters for each QoS Flow". The QoS characteristics are indication to the underlying network (and base stations) about the required performance behaviour. Within 3GPP it is not specified how this network should achieve this. The QoS Characteristics include the following:
    - Resource Type, i.e. GBR, delay critical GBR, or non GBR;
    - Priority Level; note that this is the priority in scheduling the resources among QoS Flows; this should not be confused with ARP;

- Packet Delay Budget, i.e. the upper bound for the time a packet may be delayed between UE and the UPF that is connected to the DN. For delay critical GBR QoS Flows a packet delayed more than this value is counted as lost if the transmitted data burst is less than the Maximum Data Burst Volume (see below). For all other flows, the delay budget shall be interpreted as a maximum delay with a confidence level of 98%;
- Packet Error Rate, i.e. the upper bound for the rate of packets lost between sender and receiver;
- Averaging Window for GBR QoS Flows, i.e. the duration over which GFBR and MFBR shall be calculated;
- Maximum Data Burst Volume for 5QIs with Packet Delay Budget <= 20ms, i.e. the largest amount of data that is required to be served in the period of the delay budget.

The standard 5QI table (copied from TS 23.501) is included below. The table largely resembles the corresponding 4G table. The 5QI values with extreme TNO low latency denoted by B, C, D, E, F, and G are new in 5G.