3GPP TS 22.261 V20.0.0 (2024-09)

Technical Specification

3rd Generation Partnership Project;
Technical Specification Group Services and System Aspects;
Service requirements for the 5G system;
Stage 1
(Release 20)

Keywords
Fifth, Generation, 5G

3GPP

Postal address

3GPP support office address
650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet
http://www.3gpp.org

Contents

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

The need to support different kinds of UEs (e.g. for the Internet of Things (IoT)), services, and technologies is driving the technology revolution to a high-performance and highly efficient 3GPP system. The drivers include IoT, Virtual Reality (VR), industrial control, ubiquitous on-demand coverage, as well as the opportunity to meet customized market needs. These drivers require enhancements to the devices, services, and technologies well established by 3GPP. The key objective with the 5G system is to be able to support new deployment scenarios to address diverse market segments.

This document compiles requirements that define a 5G system.

The 5G system is characterised, for example, by:

- Support for multiple access technologies
- Scalable and customizable network
- Advanced Key Performance Indicators (KPIs) (e.g. availability, latency, reliability, user experienced data rates, area traffic capacity)
- Flexibility and programmability (e.g. network slicing, diverse mobility management, Network Function Virtualization)
- Resource efficiency (both user plane and control plane)
- Seamless mobility in densely populated and heterogeneous environment
- Support for real time and non-real time multimedia services and applications with advanced Quality of Experience (QoE)

1 Scope

The present document describes the service and operational requirements for a 5G system, including a UE, NG-RAN, and 5G Core network. Requirements for a 5G E-UTRA-NR Dual Connectivity in E-UTRAN connected to EPC are found in TS 22.278 [5].

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] NGMN 5G White Paper v1.0, February 2015.

[3] 3GPP TS 22.011: "Service accessibility".

[4] Void

[5] 3GPP TS 22.278: "Service requirements for the Evolved Packet System (EPS)".

[6] 3GPP TS 22.101: "Service aspects; Service principles".

[7] 3GPP TS 22.146: "Multimedia Broadcast/Multicast Service (MBMS)".

[8] 3GPP TS 22.246: "Multimedia Broadcast/Multicast Service (MBMS) user services".

[9] 3GPP TS 22.186: "Enhancement of 3GPP support for V2X scenarios".

[10] NGMN, "Recommendations for NGMN KPIs and Requirements for 5G", June 2016

[11] 3GPP TS 22.115: "Service aspects; Charging and billing".

[12] Void

[13] Soriano, R., Alberto, M., Collazo, J., Gonzales, I., Kupzo, F., Moreno, L., & Lorenzo, J. OpenNode. Open Architecture for Secondary Nodes of the Electricity Smartgrid. In Proceedings CIRED 2011 21st International Conference on Electricity Distribution, CD1. June 2011.

[14] North American Electric Reliability Council. Frequently Asked Questions (FAQs) Cyber Security Standards CIP–002–1 through CIP–009–1. Available: http://www.nerc.com/docs/standards/sar/Revised_CIP-002-009_FAQs_06Mar06.pdf. 2006.

[15] McTaggart, Craig, et al. "Improvements in power system integrity protection schemes". Developments in Power System Protection (DPSP 2010). Managing the Change, 10th IET International Conference on. IET, 2010.

[16] IEEE Power Engineering Society – Power System Relaying Committee – System Protection Subcommittee Working Group C-6. Wide Area Protection and Emergency Control.

[17] Begovic, Miroslav, et al. "Wide-area protection and emergency control". Proceedings of the IEEE 93.5, pp. 876-891, 2005.

[18] ITU-T Recommendation G.1000 "Communications quality of service: A framework and definitions".

[19] IEC 61907, "Communication network dependability engineering".

[20] NIST, "Framework for Cyber-Physical Systems", 2016.

[21] 3GPP TS 22.104: "Service requirements for cyber-physical control applications in vertical domains".

[22] 3GPP TS 22.262: "Message Service within the 5G System".

[23] 3GPP TS 22.289: "Mobile Communication System for Railways".

[24] 3GPP TS 22.071: "Location Services".

[25] 3GPP TS 23.122: "Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode".

[26] 3GPP TS 22.125: "Unmanned Aerial System (UAS) support in 3GPP".

[27] Void

[28] 3GPP TS 22.263: "Service requirements for Video, Imaging and Audio for Professional Applications (VIAPA)".

[29] Void

[30] 3GPP TS 22.179: "Mission Critical Push to Talk (MCPTT)".

[31] IEEE 1588-2019, IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems.

[32] IEC 61850-9-3-2016 - IEC/IEEE International Standard - Communication networks and systems for power utility automation – Part 9-3: Precision time protocol profile for power utility automation.

[33] 3GPP TS 38.305: "NG Radio Access Network (NG-RAN); Stage 2 functional specification of User Equipment (UE) positioning in NG-RAN"

[34] ATIS-0900005: "Technical Report on GPS Vulnerability", https://access.atis.org/apps/group_public/download.php/36304/ATIS-0900005.pdf

[35] European Commission, Regulatory Technical Standard 25. Level of accuracy of business clocks https://ec.europa.eu/finance/securities/docs/isd/mifid/rts/160607-rts-25_en.pdf (annex https://ec.europa.eu/finance/securities/docs/isd/mifid/rts/160607-rts-25-annex_en.pdf)

[36] 5G-ACIA, "Exposure of 5G capabilities for Connected Industries and Automation Applications", 5G-ACIA white pater, February 2021, https://5g-acia.org/wp-content/uploads/2021/04/5G-ACIA_ExposureOf5GCapabilitiesForConnectedIndustriesAndAutomationApplications.pdf

[37] 3GPP TS 22.173: "IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services".

[38] ITU-T, "Technology Watch Report: The Tactile Internet", August 2014.

[39] D. Soldani, Y. Guo, B. Barani, P. Mogensen, I. Chih-Lin, S. Das, "5G for ultra-reliable low-latency communications". IEEE Network. 2018 Apr 2; 32(2):6-7.

[40] O. Holland et al., "The IEEE 1918.1 "Tactile Internet" Standards Working Group and its Standards," Proceedings of the IEEE, vol. 107, no. 2, Feb. 2019.

[41] Altinsoy, M. E., Blauert, J., & Treier, C., "Inter-Modal Effects of Non-Simultaneous Stimulus Presentation," A. Alippi (Ed.), Proceedings of the 7th International Congress on Acoustics, Rome, Italy, 2001.

[42] Hirsh I.J., and Sherrrick C.E, 1961. J. Exp. Psychol 62, 423-432

[43] Altinsoy, M.E. (2012). "The Quality of Auditory-Tactile Virtual Environments," Journal of the Audio Engineering Society, Vol. 60, No. 1/2, pp. 38-46, Jan.-Feb. 2012.

[44] M. Di Luca and A. Mahnan, "Perceptual Limits of Visual-Haptic Simultaneity in Virtual Reality Interactions," 2019 IEEE World Haptics Conference (WHC), 2019, pp. 67-72, doi: 10.1109/WHC.2019.8816173.

[45] K. Antonakoglou et al., “Toward Haptic Communications Over the 5G Tactile Internet”, IEEE Communications Surveys & Tutorials, 20 (4), 2018.

[46] ETSI GS OEU 020 (v1.1.1): "Operational energy Efficiency for Users (OEU); Carbon equivalent Intensity measurement; Operational infrastructures; Global KPIs; Global KPIs for ICT Sites".

[47] 3GPP TS 28.310: "Management and orchestration; Energy efficiency of 5G".

[48] ETSI EN 303 472: "Environmental Engineering (EE); Energy Efficiency measurement methodology and metrics for RAN equipment".

[49] 3GPP TS 32.299: " Telecommunication management; Charging management; Diameter charging applications".

[50] N. Nonaka et al., "Experimental Trial aboard Shinkansen Test Train Running at 360 km/h for 5G Evolution," 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 2022.

[51] 3GPP TS 22.137: " Service requirements for Integrated Sensing and Communication ".

[52] 3GPP TS 22.369: " Service requirements for Ambient power-enabled IoT ".

[53] 3GPP TS 22.156: "Mobile Metaverse Services”.

[54] IEEE Std 802.1Q: "IEEE Standard for Local and Metropolitan Area Networks---Bridges and Bridged Networks".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

5G enhanced positioning area: a subset of the 5G positioning service area that is assumed to be provided with additional infrastructure or deploy a particular set of positioning technologies to enhance positioning services.

NOTE 1: The enhanced positioning service area represents for example a factory plant, a dense urban area, an area along a road or railway track, a tunnel and covers both indoor and outdoor environments.

5G LAN-type service: a service over the 5G system offering private communication using IP and/or non-, i.e. UEs that are members of the same 5G LAN-VN IP type communications.

5G LAN-virtual network: a virtual network capable of supporting 5G LAN-type service.

5G satellite access network: 5G access network using at least one satellite.

5G positioning service area: a service area where positioning services would solely rely on infrastructures and positioning technologies that can be assumed to be present anywhere where 5G is present (e.g. a country-wide operator-supplied 5G network, GNSS, position/motion sensors).

NOTE 2: This includes both indoor and any outdoor environments.

active communication: a UE is in active communication when it has one or more connections established. A UE may have any combination of PS connections (e.g. PDP contexts, active PDN connections).

activity factor: percentage value of the amount of simultaneous active UEs to the total number of UEs where active means the UEs are exchanging data with the network.

Ambient IoT device: An ambient power-enabled Internet of Things device is an IoT device powered by energy harvesting, being either battery-less or with limited energy storage capability (e.g., using a capacitor).

aggregated QoS: QoS requirement(s) that apply to the traffic of a group of UEs.

area traffic capacity: total traffic throughput served per geographic area.

authorised administrator: a user or other entity authorised to partially configure and manage a network node in a CPN (e.g. a PRAS, or eRG) or a PIN element in a PIN.

carbon emission: quantity of equivalent carbon dioxide emitted (e.g. kg of CO_2 equivalent).

communication service availability: percentage value of the amount of time the end-to-end communication service is delivered according to a specified QoS, divided by the amount of time the system is expected to deliver the end-to-end service.

NOTE 3: The end point in "end-to-end" is the communication service interface.

NOTE 4: The communication service is considered unavailable if it does not meet the pertinent QoS requirements. For example, the communication service is unavailable if a message is not correctly received within a specified time, which is the sum of maximum allowed end-to-end latency and survival time.

Customer Premises Network: a network located within a premise (e.g. a residence, office or shop), which is owned, installed and/or (at least partially) configured by the customer of a public network operator.

direct device connection: the connection between two UEs without any network entity in the middle.

direct network connection: one mode of network connection, where there is no relay UE between a UE and the 5G network.

Disaster Condition: This is the condition that a government decides when to initiate and terminate, e.g. a natural disaster. When this condition applies, users may have the opportunity to mitigate service interruptions and failures.

Disaster Inbound Roamer: A user that (a) cannot get service from the PLMN it would normally be served by, due to failure of service during a Disaster Condition, and (b) is able to register with other PLMNs.

Disaster Roaming: This is the special roaming policy that applies during a Disaster Condition.

DualSteer device: A device supporting traffic steering and switching of user data (for different services) across two 3GPP access networks; it can be a single UE, in case of non-simultaneous data transmission over the two networks, or two separate UEs in case of simultaneous data transmission over the two networks.

end-to-end latency: the time that it takes to transfer a given piece of information from a source to a destination, measured at the communication interface, from the moment it is transmitted by the source to the moment it is successfully received at the destination.

energy charging rate: a means of determining the energy consumption consequence (use of energy credit) associated with charging events.

energy credit: a quantity of energy credit associated with the subscriber that can be used for credit control by the 5G system.

energy state: state of a cell, a network element and/or a network function with respect to energy, e.g. (not) energy saving states, which are defined in TS 28.310 [47].

evolved Residential Gateway: a gateway between the public operator network (fixed/mobile/cable) and a customer premises network.

holdover: A clock A, previously synchronized/syntonized to another clock B (normally a primary reference or a Master Clock) but whose frequency is determined in part using data acquired while it was synchronized/syntonized to B, is said to be in holdover or in the holdover mode as long as it is within its accuracy requirements.

NOTE 4a: holdover is defined in [31]

Holdover time: the time period that is available to repair the first priority timing source when it is lost (e.g., when the primary GNSS reference is lost). During this period the synchronization accuracy requirement should be guaranteed, e.g., by means of defining multiple synchronization references.

Hosted Service: a service containing the operator's own application(s) and/or trusted third-party application(s) in the Service Hosting Environment, which can be accessed by the user.

Hosting NG-RAN Operator: the operator that has operational control of a Shared NG-RAN.

NOTE 4b: Hosting NG-RAN Operator is a Hosting RAN Operator.

Hosting RAN Operator: as defined in 3GPP TS 22.101 [6].

hybrid access: access consisting of multiple different access types combined, such as fixed wireless access and wireline access.

indirect network connection: one mode of network connection, where there is a relay UE between a UE and the 5G network.

Indirect Network Sharing: a type of NG-RAN Sharing in which the communication between the Shared NG-RAN and the Participating Operator's core network is routed through the Hosting NG-RAN Operator's core network.

IoT device: a type of UE which is dedicated for a set of specific use cases or services and which is allowed to make use of certain features restricted to this type of UEs.

NOTE 5: An IoT device may be optimized for the specific needs of services and application being executed (e.g. smart home/city, smart utilities, e-Health and smart wearables). Some IoT devices are not intended for human type communications.

maximum energy consumption: a policy establishing an upper bound on the quantity of energy consumption [47] by the 5G system in a specific period of time, or space, e.g. energy consumption inside a given service area.

maximum energy credit limit: a policy establishing an upper bound on the aggregate quantity of energy consumption by the 5G system to provide services to a specific subscriber, e.g. in kilowatt hours.

NOTE: The terms maximum energy credit limit is distinct from 'maximum energy consumption' because the credit limit is a total amount of energy consumed, where maximum energy consumption is a limit to the consumption in a given interval of time.

network slice: a set of network functions and corresponding resources necessary to provide the required telecommunication services and network capabilities.

NG-RAN: a radio access network connecting to the 5G core network which uses NR, E-UTRA, or both.

NG-RAN Sharing: the sharing of NG-RAN among a number of operators.

non-public network: a network that is intended for non-public use.

NR: the new 5G radio access technology.

Participating NG-RAN Operator: authorized operator that is using Shared NG-RAN resources provided by a Hosting NG-RAN Operator.

NOTE 5a: Participating NG-RAN Operator is a Participating Operator.

Participating Operator: as defined in 3GPP TS 22.101 [6].

Personal IoT Network: A configured and managed group of at least one UE PIN Element and one or more PIN Element that communicate with each other.

PIN Element: UE or non-3GPP device that can communicate within a PIN.

PIN direct connection: the connection between two PIN Elements without any 3GPP RAN or core network entity in the middle.

NOTE 5A: A PIN direct connection could internally be relayed by other PIN Elements.

NOTE 5B: When a PIN direct connection is between two PIN Elements that are UEs this direct connection is typically known as a direct device connection.

PIN Element with Gateway Capability: a UE PIN Element that has the ability to provide connectivity to and from the 5G network for other PIN Elements.

NOTE 5C: A PIN Element can have both PIN management capability and Gateway Capability.

PIN Element with Management Capability: A PIN Element with capability to manage the PIN.

positioning service availability: percentage value of the amount of time the positioning service is delivering the required position-related data within the performance requirements, divided by the amount of time the system is expected to deliver the positioning service according to the specification in the targeted service area.

proximity-based work task offloading: a relay UE receives data from a remote UE via direct device connection and performs calculation of a work task for the remote UE. The calculation result can be further sent to network server.

positioning service latency: time elapsed between the event that triggers the determination of the position-related data and the availability of the position-related data at the system interface.

Premises Radio Access Station: a base station installed at a customer premises network.

priority service: a service that requires priority treatment based on regional/national or operator policies.

private communication: a communication between two or more UEs belonging to a restricted set of UEs**.**

private network: an isolated network deployment that does not interact with a public network.

private slice: a dedicated network slice deployment for the sole use by a specific third-party.

ProSe UE-to-UE Relay: a Public Safety ProSe-enabled UE that acts as a relay between two other Public Safety ProSe-enabled UEs.

Ranging: refers to the determination of the distance between two UEs and/or the direction of one UE from the other one via direct device connection.

relative positioning: relative positioning is to estimate position relatively to other network elements or relatively to other UEs.

reliability: in the context of network layer packet transmissions, percentage value of the packets successfully delivered to a given system entity within the time constraint required by the targeted service out of all the packets transmitted.

renewable energy: energy from renewable sources as energy from renewable non-fossil sources, namely wind, solar, aerothermal, geothermal, hydrothermal and ocean energy, hydropower, biomass, landfill gas, sewage treatment plant gas and biogases

NOTE 2: This definition was taken from [48].

satellite: a space-borne vehicle embarking a bent pipe payload or a regenerative payload telecommunication transmitter, placed into Low-Earth Orbit (LEO) typically at an altitude between 300 km to 2 000 km, Medium-Earth Orbit (MEO) typically at an altitude between 8 000 to 20 000 k m, or Geostationary satellite Earth Orbit (GEO) at 35 786 km altitude.

satellite access: direct connectivity between the UE and the satellite.

satellite NG-RAN: a NG-RAN which uses NR in providing satellite access to UEs.

service area: geographic region where a 3GPP communication service is accessible.

NOTE 6: The service area can be indoors.

NOTE 7: For some deployments, e.g. in process industry, the vertical dimension of the service area can be considerable.

service continuity: the uninterrupted user experience of a service that is using an active communication when a UE undergoes an access change without, as far as possible, the user noticing the change.

NOTE 8: In particular service continuity encompasses the possibility that after a change the user experience is maintained by a different telecommunication service (e.g. tele- or bearer service) than before the change.

NOTE 9: Examples of access changes include the following. For EPS: CS/PS domain change. For EPS and 5G: radio access change, switching between a direct network connection and an indirect network connection.

Service Hosting Environment: the environment, located inside of 5G network and fully controlled by the operator, where Hosted Services are offered from.

serving satellite: a satellite providing the satellite access to an UE. In the case of NGSO, the serving satellite is always changing due to the nature of the satellite constellation.

Shared NG-RAN: as defined in 3GPP TS 22.101 [6].

NOTE: Shared NG-RAN can be a shared satellite NG-RAN.

Stand-alone Non-Public Network: A non-public network not relying on network functions provided by a PLMN

SNPN Credential Provider: Entity within the 5G system that creates and manages identity information and provides authentication services for those identities for the purpose of accessing a SNPN

NOTE: The SNPN Credential Provider can also authorize access to a non-public network for a subscriber associated with an identity handled by this SNPN Credential Provider.

S&F Satellite operation: S&F (Store and Forward) Satellite operation is an operation mode of a 5G system with satellite-access where the 5G system can provide some level of service (in storing and forwarding the data) when satellite connectivity is intermittently/temporarily unavailable, e.g. to provide communication service for UEs under satellite coverage without a simultaneous active feeder link connection to the ground segment.

S&F data retention period: it is the data storage validity period for a 5G system with satellite access supporting store and forward operation (e.g. after which undelivered data stored is being discarded).

synchronization threshold: A synchronization threshold can be defined as the maximum tolerable temporal separation of the onset of two stimuli, one of which is presented to one sense and the other to another sense, such that the accompanying sensory objects are perceived as being synchronous.

NOTE 10: This definition is based on [41].

survival time: the time that an application consuming a communication service may continue without an anticipated message.

Time to First Fix (TTFF): time elapsed between the event triggering for the first time the determination of the position-related data and the availability of the position-related data at the positioning system interface.

Traffic steering: the procedure that selects an access network and transfers traffic over the selected access network. This can apply to traffic of one or multiple services/applications across two 3GPP access networks, including scenarios where all services use the same network connection (no simultaneous data over the two networks) or different services are steered across different networks (with simultaneous data over the two networks).

Traffic switching: the procedure that moves all traffic from one access network to another access network in a way that minimizes service interruption. This can apply to traffic of one or multiple services/applications across two 3GPP access networks, including scenarios where all services use the same network connection (no simultaneous data over the two networks) or different services are moved to different networks (with simultaneous data over the two networks).

UE-Satellite-UE communication: for a 5G system with satellite access, it refers to the communication between UEs under the coverage of one or more serving satellites, using satellite access without the user traffic going through the ground segment.

User Equipment: An equipment that allows a user access to network services via 3GPP and/or non-3GPP accesses.

user experienced data rate: the minimum data rate required to achieve a sufficient quality experience, with the exception of scenario for broadcast like services where the given value is the maximum that is needed.

wireless backhaul: a link which provides an interconnection between 5G network nodes and/or transport network using 5G radio access technology**.**

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

5G LAN-VN	5G LAN-Virtual Network
A/S	Actuator/Sensor
CPN	Customer Premises Network
eFMSS	Enhancement to Flexible Mobile Service Steering
eRG	Evolved Residential Gateway
eV2X	Enhanced V2X
FL	Federated Learning
FMSS	Flexible Mobile Service Steering
GEO	Geostationary satellite Earth Orbit
ICP	Internet Content Provider
ID	Identification
IMU	Inertial Measurement Unit
IOPS	Isolated E-UTRAN Operation for Public Safety
ISL	Inter-Satellite Link
LEO	Low-Earth Orbit
MBS	Metropolitan Beacon System
MCS	Mission Critical Services
MCX	Mission Critical X, with X = PTT or X = Video or X = Data
MEO	Medium-Earth Orbit
MIoT	Massive Internet of Things
MMTEL	Multimedia Telephony
MOCN	Multi-Operator Core Network
MPS	Multimedia Priority Service
MSGin5G	Message Service Within the 5G System
MSTP	Multiple Spanning Tree Protocol
NGSO	Non-Geostationary Satellite Orbit
NPN	Non-Public Network
PIN	Personal IoT Network
PRAS	Premises Radio Access Station
RSTP	Rapid Spanning Tree Protocol
RVAS	Roaming Value-Added Service
SEES	Service Exposure and Enablement Support
SNPN	Stand-alone Non-Public Network
S&F	Store and Forward
SST	Slice/Service Type
TBS	Terrestrial Beacon System
TTFF	Time To First Fix
UAV	Uncrewed Aerial Vehicle
UTC	Coordinated Universal Time
XR	eXtended Reality

4 Overview

Unlike previous 3GPP systems that attempted to provide a 'one size fits all' system, the 5G system is expected to be able to provide optimized support for a variety of different services, different traffic loads, and different end user communities. Various industry white papers, most notably, the NGMN 5G White Paper [2], describe a multi-faceted 5G system capable of simultaneously supporting multiple combinations of reliability, latency, throughput, positioning, and availability. This technology revolution is achievable with the introduction of new technologies, both in access and the core, such as flexible, scalable assignment of network resources. In addition to increased flexibility and optimization, a 5G system needs to support stringent KPIs for latency, reliability, throughput, etc. Enhancements in the radio interface contribute to meeting these KPIs as do enhancements in the core network, such as network slicing, in-network caching and hosting services closer to the end points.

A 5G system also supports new business models such as those for IoT and enterprise managed networks. Drivers for the 5G KPIs include services such as Uncrewed Aerial Vehicle (UAV) control, Augmented Reality (AR), and factory

automation. Network flexibility enhancements support self-contained enterprise networks, installed and maintained by network operators while being managed by the enterprise. Enhanced connection modes and evolved security facilitate support of massive IoT, expected to include tens of millions of UEs sending and receiving data over the 5G network.

Flexible network operations are the mainstay of the 5G system. The capabilities to provide this flexibility include network slicing, network capability exposure, scalability, and diverse mobility. Other network operations requirements address the necessary control and data plane resource efficiencies, as well as network configurations that optimize service delivery by minimizing routing between end users and application servers. Enhanced charging and security mechanisms handle new types of UEs connecting to the network in different ways. The enhanced flexibility of the 5G system also allows to cater to the needs of various verticals. For example, the 5G system introduces the concept of non-public networks providing exclusive access for a specific set of users and specific purpose(s). Non-public networks can, depending on deployment and (national) regulations, support different subsets of 5G functionality. In this specification 5G network requirements apply to both NPNs and PLMNs, unless specified otherwise. Additionally, there are specific requirements dedicated only to NPNs or PLMNs, which are indicated accordingly. More information can be found in Section 6.25.

Mobile Broadband (MBB) enhancements aim to meet a number of new KPIs. These pertain to high data rates, high user density, high user mobility, highly variable data rates, deployment, and coverage. High data rates are driven by the increasing use of data for services such as streaming (e.g. video, music, and user generated content), interactive services (e.g. AR), and IoT. These services come with stringent requirements for user experienced data rates as well as associated requirements for latency to meet service requirements. Additionally, increased coverage in densely populated areas such as sports arenas, urban areas, and transportation hubs has become essential for pedestrians and users in urban vehicles. New KPIs on traffic and connection density enable both the transport of high volumes of data traffic per area (traffic density) and transport of data for a high number of connections (e.g. UE density or connection density). Many UEs are expected to support a variety of services which exchange either a very large (e.g. streaming video) or very small (e.g. data burst) amount of data. The 5G system will handle this variability in a resource efficient manner. All of these cases introduce new deployment requirements for indoor and outdoor, local area connectivity, high user density, wide area connectivity, and UEs travelling at high speeds.

Another aspect of 5G KPIs includes requirements for various combinations of latency and reliability, as well as higher accuracy for positioning. These KPIs are driven by support for both commercial and public safety services. On the commercial side, industrial control, industrial automation, UAV control, and AR are examples of those services. Services such as UAV control will require more precise positioning information that includes altitude, speed, and direction, in addition to horizontal coordinates.

Support for Massive Internet of Things (MIoT) brings many new requirements in addition to those for the enhanced KPIs. The expansion of connected things introduces a need for significant improvements in resource efficiency in all system components (e.g. UEs, IoT devices, radio, access network, core network).

The 5G system also aims to enhance its capability to meet KPIs that emerging V2X applications require. For these advanced applications, the requirements, such as data rate, reliability, latency, communication range and speed, are made more stringent.

5 High-level requirements

5.1 Migration to 5G

5.1.1 Description

The 5G system supports most of the existing EPS services, in addition to many new services. The existing EPS services can be accessed using the new 5G access technologies even where the EPS specifications might indicate E-UTRA(N) only. Only new or changed service requirements for new or changed services are specified in this TS. The few EPS capabilities that are not supported by the 5G system are identified in clause 5.1.2.2 below.

5.1.2 Requirements

5.1.2.1 Interworking between 5G systems

The 5G system shall support a UE with a 5G subscription roaming into a 5G Visited Mobile Network which has a roaming agreement with the UE's 5G Home Mobile Network.

The 5G system shall enable a Visited Mobile Network to provide support for establishing home network provided data connectivity as well as visited network provided data connectivity.

The 5G system shall enable a Visited Mobile Network to provide support for services provided in the home network as well as provide services in the visited network. Whether a service is provided in the visited network or in the home network is determined on a service by service basis.

The 5G system shall provide a mechanism for a network operator to limit access to its services for a roaming UE, (e.g. based on roaming agreement).

The 5G system shall provide a mechanism for a network operator to direct a UE onto a partnership network for routing all or some of the UE user plane and associated control plane traffic over the partnership network, subject to an agreement between the operators.

5.1.2.2 Legacy service support

In principle, the 5G system shall support all EPS capabilities (e.g. from TSs 22.011, 22.101, 22.278, 22.185, 22.071, 22.115, 22.153, 22.173, 22.468). However,

- voice service continuity from NG-RAN to GERAN shall not be supported,
- voice service continuity from NG-RAN to UTRAN CS should be supported (see Note),
- voice service continuity from GERAN to NG-RAN shall not be supported,
- voice service continuity from UTRAN to NG-RAN shall not be supported,
- CS fallback from NG-RAN to GERAN shall not be supported,
- CS fallback from NG-RAN to UTRAN shall not be supported,
- seamless handover between NG-RAN and GERAN shall not be supported,
- seamless handover between NG-RAN and UTRAN shall not be supported,
- access to a 5G core network via GERAN or UTRAN shall not be supported,
- video service continuity between 5GS and UMTS shall not be supported,
- IP address preservation for PS service when UE moves between 5GS and GSM/UMTS shall not be supported, and
- Service continuity between 5GS and CDMA2000 shall not be supported.

NOTE: Architectural or protocol changes needed to support voice service continuity from NG-RAN to UTRAN CS are expected to have minimum impact on architecture, specifications, or the development of the 5G New Core and New Radio.

5.1.2.3 Interoperability with legacy 3GPP systems

The 5G system shall support mobility procedures between a 5G core network and an EPC with minimum impact to the user experience (e.g. QoS, QoE).

6 Basic capabilities

6.1 Network slicing

6.1.1 Description

Network slicing allows the operator to provide customised networks. For example, there can be different requirements on functionality (e.g. priority, charging, policy control, security, and mobility), differences in performance requirements (e.g. latency, mobility, availability, reliability and data rates), or they can serve only specific users (e.g. MPS users, Public Safety users, corporate customers, roamers, or hosting an MVNO).

A network slice can provide the functionality of a complete network, including radio access network functions, core network functions (e.g. potentially from different vendors) and IMS functions. One network can support one or several network slices.

6.1.2 Requirements

6.1.2.1 General

The serving 5G network shall support providing connectivity to home and roaming users in the same network slice.

In shared 5G network configuration, each operator shall be able to apply all the requirements from this clause to their allocated network resources.

The 5G system shall be able to support IMS as part of a network slice.

The 5G system shall be able to support IMS independent of network slices.

For a UE authorized to access multiple network slices of one operator which cannot be simultaneously used by the UE (e.g. due to radio frequency restrictions), the 5G system shall be able to support the UE to access the most suitable network slice in minimum time (e.g. based on the location of the UE, ongoing applications, UE capability, frequency configured for the network slice).

5G system shall minimize signalling exchange and service interruption time for a network slice, e.g. when restrictions related to radio resources change (e.g., frequencies, RATs).

For a roaming UE activating a service/application requiring a network slice not offered by the serving network but available in the area from other network(s), the HPLMN shall be able to provide the UE with prioritization information of the VPLMNs with which the UE may register for the network slice.

The 5G system shall be able to minimize power consumption of a UE (e.g. reduce unnecessary cell measurements), in an area where no authorized network slice is available.

When a UE moves out of the service area of a network slice for an active application, the 5G system shall be able to minimize impact on the active applications (e.g., providing early notification).

NOTE 1: Various methods can be used to detect whether the UE moves toward the border area and to notify the UE.

The 5G system shall support a mechanism for a UE to select and access network slice(s) based on UE capability, ongoing application, radio resources assigned to the slice, and policy (e.g., application preference).

The 5G system shall support a mechanism to optimize resources of network slices (e.g., due to operator deploying different frequency to offer different network slices) based on network slice usage patterns and policy (e.g., application preference) of a UE or group of UEs

For UEs that have the ability to obtain service from more than one VPLMN simultaneously, the following requirements apply:

- When a roaming UE with a single PLMN subscription requires simultaneous access to multiple network slices and the network slices are not available in a single VPLMN, the 5G system shall enable the UE to:

- be registered to more than one VPLMN simultaneously; and

- use network slices from more than one VPLMN simultaneously

- The HPLMN shall be able to authorise a roaming UE with a single PLMN subscription to be registered to more than one VPLMN simultaneously in order to access network slices of those VPLMNs.

- The HPLMN shall be able to provide a UE with permission and prioritisation information of the VPLMNs the UE is authorised to register to in order to use specific network slices.

NOTE 2: The above requirements assume certain UE capabilities, e.g. the ability to be connected to more than one PLMN simultaneously.

6.1.2.2 Management

The 5G system shall allow the operator to create, modify, and delete a network slice.

The 5G system shall allow the operator to define and update the set of services and capabilities supported in a network slice.

The 5G system shall allow the operator to configure the information which associates a UE to a network slice.

The 5G system shall allow the operator to configure the information which associates a service to a network slice.

The 5G system shall allow the operator to assign a UE to a network slice, to move a UE from one network slice to another, and to remove a UE from a network slice based on subscription, UE capabilities, the access technology being used by the UE, operator's policies and services provided by the network slice.

The 5G system shall support a mechanism for the VPLMN, as authorized by the HPLMN, to assign a UE to a network slice with the needed services or to a default network slice.

The 5G system shall enable a UE to be simultaneously assigned to and access services from more than one network slice of one operator.

Traffic and services in one network slice shall have no impact on traffic and services in other network slices in the same network.

Creation, modification, and deletion of a network slice shall have no or minimal impact on traffic and services in other network slices in the same network.

The 5G system shall support scaling of a network slice, i.e. adaptation of its capacity.

The 5G system shall enable the network operator to define a minimum available capacity for a network slice. Scaling of other network slices on the same network shall have no impact on the availability of the minimum capacity for that network slice.

The 5G system shall enable the network operator to define a maximum capacity (e.g., number of UEs, number of data sessions) for a network slice.

The 5G system shall enable the network operator to define a priority order between different network slices in case multiple network slices compete for resources on the same network.

The 5G system shall support means by which the operator can differentiate policy control, functionality and performance provided in different network slices.

6.1.2.3 Network slice constraints

The 5G system shall support a mechanism to prevent a UE from trying to access a radio resource dedicated to a specific private slice for any purpose other than that authorized by the associated third-party.

NOTE 1: UEs that are not authorized to access a specific private slice will not be able to access it for emergency calls if the private slice does not support emergency services.

The 5G system shall support a mechanism to configure a specific geographic area in which a network slice is accessible, i.e. a UE shall be within the geographical area in order to access the network slice.

The 5G system shall support a mechanism to limit a UE to only receiving service from an authorized slice.

For a UE authorized to access to multiple network slices of one operator which cannot be simultaneously used by the UE (e.g. due to radio frequency restrictions), the 5G system shall minimize service interruption time when the UE changes the access from one network slice to another network slice. (e.g. based on changes of active applications).

For traffic pertaining to a network slice offered via a relay node, 5G system shall use only radio resources (e.g. frequency band) allowed for the network slice.

NOTE 2: Allowed radio resources (e.g., frequency band) may be different for direct network connections (between UE and NG-RAN) than for backhaul connections (between the relay node and the NG-RAN).

The 5G system shall support a mechanism to prevent a UE from registering with a network slice when the maximum number of UEs for that slice are registered.

The 5G system shall support a mechanism to prevent a UE from establishing a new data session within a network slice when the maximum number of data sessions for that slice are established.

NOTE 3: Based on national/regional regulations and operator policy, exemptions can be provided for UEs configured for priority services (e.g., MPS) and for priority service sessions.

6.1.2.4 Cross-network slice coordination

The 5G system shall support a mechanism to provide time stamps with a common time base at the monitoring API, for services that cross multiple network slices and 5G networks.

The 5G system shall provide suitable APIs to coordinate network slices in multiple 5G networks so that the selected communication services of a non-public network can be extended through a PLMN (e.g. the service is supported by a slice in the non-public network and a slice in the PLMN).

The 5G system shall provide a mechanism to enable an MNO to operate a hosted non-public network and private slice(s) of its PLMN associated with the hosted non-public network in a combined manner.

6.2 Diverse mobility management

6.2.1 Description

A key feature of 5G is support for UEs with different mobility management needs. 5G will support UEs with a range of mobility management needs, including UEs that are

- stationary during their entire usable life (e.g. sensors embedded in infrastructure),
- stationary during active periods, but nomadic between activations (e.g. fixed access),
- mobile within a constrained and well-defined space (e.g. in a factory), and
- fully mobile.

Moreover, some applications require the network to ensure seamless mobility of a UE so that mobility is hidden from the application layer to avoid interruptions in service delivery while other applications have application specific means to ensure service continuity. But these other applications can still require the network to minimize interruption time to ensure that their application-specific means to ensure service continuity work effectively.

With the ever-increasing multimedia broadband data volumes, it is also important to enable the offloading of IP traffic from the 5G network onto traditional IP routing networks via an IP anchor node close to the network edge. As the UE moves, changing the IP anchor node can be needed in order to reduce the traffic load in the system, reduce end-to-end latency and provide a better user experience.

The flexible nature of a 5G system will support different mobility management methods that minimize signalling overhead and optimize access for these different types of UEs.

6.2.2 General requirements

The 5G network shall allow operators to optimize network behaviour (e.g. mobility management support) based on the mobility patterns (e.g. stationary, nomadic, spatially restricted mobility, full mobility) of a UE or group of UEs.

The 5G system shall enable operators to specify and modify the types of mobility support provided for a UE or group of UEs.

The 5G system shall optimize mobility management support for a UE or group of UEs that use only mobile originated communications.

The 5G system shall support inter- and/or intra- access technology mobility procedures within 5GS with minimum impact to the user experience (e.g. QoS, QoE).

6.2.3 Service continuity requirements

The 5G system shall enable packet loss to be minimized during inter- and/or intra- access technology changes for some or all connections associated with a UE.

The 5G system shall minimize interruption time during inter- and/or intra- access technology mobility for some or all connections associated with a UE.

NOTE: The interruption time includes all delays which have impact on service continuity.

For applications that require the same IP address during the lifetime of the session, the 5G system shall enable maintaining the IP address assigned to a UE when moving across different cells and access technologies for connections associated with a UE.

The 5G system shall enable minimizing impact to the user experience (e.g. minimization of interruption time) when changing the IP address and IP anchoring point for some or all connections associated with a UE.

The 5G system shall support service continuity for a remote UE, when the remote UE changes from a direct network connection to an indirect network connection and vice-versa.

The 5G system shall support service continuity for a remote UE, when the remote UE changes from one relay UE to another and both relay UEs use 3GPP access to the 5G core network.

Satellite access related service continuity requirements are covered in clause 6.46.3.

6.2.4 Roaming related requirements

Satellite access related roaming requirements are covered in clause 6.46.4.

6.3 Multiple access technologies

6.3.1 Description

The 5G system will support 3GPP access technologies, including one or more NR and E-UTRA as well as non-3GPP access technologies. Interoperability among the various access technologies will be imperative. For optimization and resource efficiency, the 5G system will select the most appropriate 3GPP or non-3GPP access technology for a service, potentially allowing multiple access technologies to be used simultaneously for one or more services active on a UE. New technology such as satellite and wide area base stations will increase coverage and availability. This clause provides requirements for interworking with the various combinations of access technologies.

6.3.2 Requirements

6.3.2.1 General

Based on operator policy, the 5G system shall enable the UE to select, manage, and efficiently provision services over the 3GPP or non-3GPP access.

Based on operator policy, the 5G system shall support steering a UE to select certain 3GPP access network(s).

Based on operator policy, the 5G system shall be able to dynamically offload part of the traffic (e.g. from 3GPP RAT to non-3GPP access technology), taking into account traffic load and traffic type.

Based on operator policy, the 5G system shall be able to provide simultaneous data transmission via different access technologies (e.g. NR, E-UTRA, non-3GPP), to access one or more 3GPP services.

When a UE is using two or more access technologies simultaneously, the 5G system shall be able to optimally distribute user traffic over the access technologies in use, taking into account e.g. service, traffic characteristics, radio characteristics, and UE's moving speed.

The 5G system shall be able to support data transmissions optimized for different access technologies (e.g. 3GPP, non-3GPP) and accesses to local data networks (e.g. local traffic routing) for UEs that are simultaneously connected to the network via different accesses.

NOTE: This applies to the scenario with simultaneous 3GPP and non-3GPP accesses.

Based on operator policy, the 5G system shall be able to add or drop the various access connections for a UE during a session.

The 5G system shall be able to support mobility between the supported access networks (e.g. NG-RAN, WLAN, fixed broadband access network, 5G satellite access network).

The 5G system shall support UEs with multiple radio and single radio capabilities.

The 5G system shall support dynamic and static network address allocation of a common network address to the UE over all supported access types.

The 5G system shall support a set of identities for a single user in order to provide a consistent set of policies and a single set of services across 3GPP and non-3GPP access types.

The 5G system shall support the capability to operate in licensed and/or unlicensed bands.

6.3.2.2 E-UTRA access

The 5G system shall be able to support seamless handover between NR and E-UTRA.

The 5G system shall support UEs with dual radio capability (i.e. a UE that can transmit on NR and E-UTRA simultaneously) as well as UEs with single radio capability (i.e. a UE that cannot transmit on NR and E-UTRA simultaneously).

6.3.2.3 Satellite access

The 5G system shall be able to provide services using satellite access.

NOTE: Additional requirements related to satellite access can be found in clause 6.46.

6.3.2.4 Fixed broadband access

The 5G system shall be able to efficiently support connectivity using fixed broadband access.

NOTE: The specification of fixed broadband access network is outside the scope of 3GPP.

The 5G system shall support use of a relay UE that supports multiple access types (e.g. 5G RAT, WLAN access, fixed broadband access).

The 5G system shall support use of a home base station that supports multiple access types (e.g. 5G RAT, WLAN access, fixed broadband access).

6.4 Resource efficiency

6.4.1 Description

5G introduces the opportunity to design a system to be optimized for supporting diverse UEs and services. While support for IoT is provided by EPS, there is room for improvement in efficient resource utilization that can be designed into a 5G system whereas they are not easily retrofitted into an existing system. Some of the underlying principles of the potential service and network operation requirements associated with efficient configuration, deployment, and use of UEs in the 5G network include bulk provisioning, resource efficient access, optimization for UE originated data transfer, and efficiencies based on the reduced needs related to mobility management for stationary UEs and UEs with restricted range of movement.

As sensors and monitoring UEs are deployed more extensively, the need to support UEs that send data packages ranging in size from a small status update in a few bits to streaming video increases. A similar need exists for smart phones with widely varying amounts of data. Specifically, to support short data bursts, the network should be able to operate in a mode where there is no need for a lengthy and high overhead signalling procedure before and after small amounts of data are sent. The system will, as a result, avoid both a negative impact to battery life for the UE and wasting signalling resources.

For small form factor UEs it will be challenging to have more than 1 antenna due to the inability to get good isolation between multiple antennas. Thus, these UEs need to meet the expected performance in a 5G network with only one antenna.

Cloud applications like cloud robotics perform computation in the network rather than in a UE, which requires the system to have high data rate in the uplink and very low round trip latency. Supposed that high density cloud robotics will be deployed in the future, the 5G system need to optimize the resource efficiency for such scenario.

Additional resource efficiencies will contribute to meeting the various KPIs defined for 5G. Control plane resource efficiencies can be achieved by optimizing and minimizing signalling overhead, particularly for small data transmissions. Mechanisms for minimizing user plane resources utilization include in-network caching and application in a Service Hosting Environment closer to the end user. These optimization efforts contribute to achieving lower latency and higher reliability.

Diverse mobility management related resource efficiencies are covered in clause 6.2.

Security related resource efficiencies are covered in clause 8.8.

6.4.2 Requirements

6.4.2.1 General

The 5G system shall minimize control and user plane resource usage for data transfer from send only UEs.

The 5G system shall minimize control and user plane resource usage for stationary UEs (e.g. lower signalling to user data resource usage ratio).

The 5G system shall minimize control and user plane resource usage for transfer of infrequent small data units.

The 5G system shall optimize the resource use of the control plane and/or user plane for transfer of small data units.

The 5G system shall optimize the resource use of the control plane and/or user plane for transfer of continuous uplink data that requires both high data rate (e.g. 10 Mbit/s) and very low end-to-end latency (e.g. 1-10 ms).

The 5G network shall optimize the resource use of the control plane and/or user plane to support high density connections (e.g. 1 million connections per square kilometre) taking into account, for example, the following criteria:

- type of mobility support;
- communication pattern (e.g. send-only, frequent or infrequent);
- characteristics of payload (e.g. small or large size data payload);
- characteristics of application (e.g. provisioning operation, normal data transfer);

- UE location;
- timing pattern of data transfer (e.g. real time or non-delay sensitive).

The 5G system shall efficiently support service discovery mechanisms where UEs can discover, subject to access rights:

- status of other UEs (e.g. sound on/off);
- capabilities of other UEs (e.g. the UE is a relay UE) and/or;
- services provided by other UEs (e.g. the UE is a colour printer).

The 5G system shall be able to minimise the amount of wireless backhaul traffic (e.g. consolidating data transmissions to 1 larger rather than many smaller), when applicable (e.g. providing service in an area subject to power outages).

The 5G system shall support small form factor UEs with single antenna.

NOTE: Small form factor UEs are typically expected to have the diagonal less than 1/5 of the lowest supported frequency wave length.

Satellite access related resource efficiency requirements are covered in clause 6.46.5.

6.4.2.2 Efficient bulk operations for IoT

The 5G network shall optimize the resource use of the control plane and/or user plane to support bulk operation for high connection density (e.g. 1 million connections per square kilometre) of multiple UEs.

The 5G system shall support a timely, efficient, and/or reliable mechanism to transmit the same information to multiple UEs.

6.4.2.3 Efficient management for IoT

The 5G network shall optimize the resource use of the control plane and/or user plane to manage (e.g. provide service parameters, activate, deactivate) a UE.

The 5G network shall be able to provide policies for background data transfer to a UE so that the 5G system can optimally use the control plane and/or user plane resources.

6.4.2.4 Efficient control plane

The 5G system shall minimize the signalling that is required prior to user data transmission.

NOTE: The amount of signalling overhead may vary based on the amount of data to be transmitted, even for the same UE.

6.5 Efficient user plane

6.5.1 Description

5G is designed to meet diverse services with different and enhanced performances (e.g. high throughput, low latency and massive connections) and data traffic model (e.g. IP data traffic, non-IP data traffic, short data bursts and high throughput data transmissions).

User plane should be more efficient for 5G to support differentiated requirements. On one hand, a Service Hosting Environment located inside of operator's network can offer Hosted Services closer to the end user to meet localization requirement like low latency, low bandwidth pressure. These Hosted Services contain applications provided by operators and/or trusted 3rd parties. On the other hand, user plane paths can be selected or changed to improve the user experience or reduce the bandwidth pressure, when a UE or application changes location during an active communication, or due to operational needs in the service hosting environment (e.g. based on usage information).

The 5G network can also support multiple wireless backhaul connections (e.g. satellites and/or terrestrial), and efficiently route and/or bundle traffic among them.

6.5.2 Requirements

Based on operator policy, application needs, or both, the 5G system shall support an efficient user plane path between UEs attached to the same network, modifying the path as needed when the UE moves during an active communication.

The 5G network shall enable a Service Hosting Environment provided by operator.

Based on operator policy, the 5G network shall be able to support routing of data traffic between a UE attached to the network and an application in a Service Hosting Environment for specific services, modifying the path as needed when the UE moves during an active communication.

Based on operator policy, application needs, or both, the 5G system shall support an efficient user plane path, modifying the path as needed when the UE moves or application changes location, between a UE in an active communication and:

- an application in a Service Hosting Environment; or
    - an application server located outside the operator's network; or
    - an application server located in a customer premises network or personal IoT network.

The 5G network shall maintain user experience (e.g. QoS, QoE) when a UE in an active communication moves from a location served by a Service Hosting Environment to:

- another location served by a different Service Hosting Environment; or
- another location served by an application server located outside the operator's network; or
- another location served by an application server located in a customer premises network or personal IoT network, and vice versa.

The 5G network shall maintain user experience (e.g. QoS, QoE) when an application for a UE moves as follows:

- within a Service Hosting Environment; or
- from a Service Hosting Environment to another Service Hosting Environment; or
- from a Service Hosting Environment to an application server located place outside the operator's network; or
- from a Service Hosting Environment to an application server located in a customer premises network or personal IoT network, and vice versa.

The 5G network shall be able to interact with applications in a Service Hosting Environment for efficient network resource utilization and offloading data traffic to the most suitable Service Hosting Environment, e.g. close to the UE's point of attachment to the access network or based on usage information.

NOTE: To accomplish offloading data traffic, usage information might be exposed to the Service Hosting Environment.

The 5G network shall support configurations of the Service Hosting Environment in the network (e.g. access network, core network), that provide application access close to the UE's point of attachment to the access network.

The 5G system shall support mechanisms to enable a UE to access the closest Service Hosting Environment for a specific hosted application or service.

The 5G network shall enable instantiation of applications for a UE in a Service Hosting Environment close to the UE's point of attachment to the access network.

The 5G system shall be able to suspend or stop application instances in a Service Hosting Environment.

NOTE: Not all applications will always be available in all Service Hosting Environments. Therefore, it may be needed to instantiate an application at a Service Hosting Environment nearby for serving a particular UE.

Based on operator policy, the 5G system shall provide a mechanism such that one type of traffic (from a specific application or service) to/from a UE can be offloaded close to the UE's point of attachment to the access network, while not impacting other traffic type to/from that same UE.

Satellite access related efficient user plane requirements are covered in clause 6.46.6.

The 5G System shall enable the discovery of a suitable Hosted Service.

6.6 Efficient content delivery

6.6.1 Description

Video-based services (e.g. live streaming, VR) and personal data storage applications have been instrumental for the massive growth in mobile broadband traffic. Subject to service agreement between the operator and the content provider, the information of content and content itself can be aware by operator. In-network content caching provided by the operator, a third-party or both, can improve user experience, reduce backhaul resource usage and utilize radio resource efficiently.

The operation of in-network caching includes flexible management of the location of the content cache within the network and efficient delivery of content to and from the appropriate content caching application. Examples of services are the delivery of popular video content from a content caching application via broadcast, and secure storage of a user's personal data or files using a distributed caching application. Such a service could also provide a student with a wireless backpack, where students can resume their work through the same or a different UE at any time, with very fast response times from the network.

6.6.2 Requirements

The 5G system shall enable efficient delivery of content from a content caching application under the control of the operator (e.g. a cache located close to the UE).

The 5G system shall support a content caching application in a UE under the control of the operator.

The 5G system shall support configurations of content caching applications in the network (e.g. access network, core network), that provide content close to the UE.

Based on operator policy, the 5G system shall support an efficient mechanism for selection of a content caching application (e.g. minimize utilization of radio, backhaul resources and/or application resource) for delivery of the cached content to the UE.

The 5G system shall support a mechanism for the operator to manage content distribution across content caching applications.

The 5G system shall support delivery of cached content from a content caching application via the broadcast/multicast service.

For a 5G system with satellite access, the following requirement applies.

- A 5G system with satellite access shall be able to optimise the delivery of content from a content caching application by taking advantage of satellites in supporting ubiquitous service, as well as broadcasting/multicasting on very large to global coverages.

6.7 Priority, QoS, and policy control

6.7.1 Description

The 5G network will support many commercial services (e.g. medical) and regional or national regulatory services (e.g. MPS, Emergency, Public Safety) with requirements for priority treatment. Some of these services share common QoS characteristics such as latency and packet loss rate but can have different priority requirements. For example, UAV control and air traffic control can have stringent latency and reliability requirements but not necessarily the same priority requirements. In addition, voice-based services for MPS and Emergency share common QoS characteristics as applicable for normal public voice communications yet can have different priority requirements. The 5G network will need to support mechanisms that enable the decoupling of the priority of a particular communication from the associated QoS characteristics such as latency and reliability to allow flexibility to support different priority services

(that need to be configurable to meet operator needs, consistent with operator policies and corresponding national and regional regulatory policies).

The network needs to support flexible means to make priority decisions based on the state of the network (e.g. during disaster events and network congestion) recognizing that the priority needs can change during a crisis. The priority of any service can be different for a user of that service based on operational needs and regional or national regulations. Therefore, the 5G system should allow a flexible means to prioritise and enforce prioritisation among the services (e.g. MPS, Emergency, medical, Public Safety) and among the users of these services. The traffic prioritisation can be enforced by adjusting resource utilization or pre-empting lower priority traffic.

The network must offer means to provide the required QoS (e.g. reliability, latency, and bandwidth) for a service and the ability to prioritize resources when necessary to meet the service requirements. Existing QoS and policy frameworks handle latency and improve reliability by traffic engineering. In order to support 5G service requirements, it is necessary for the 5G network to offer QoS and policy control for reliable communication with latency required for a service and enable the resource adaptations as necessary.

The network needs to allow multiple services to coexist, including multiple priority services (e.g. Emergency, MPS and MCS) and must provide means to prevent a single service from consuming or monopolizing all available network resources, or impacting the QoS (e.g. availability) of other services competing for resources on the same network under specific network conditions. For example, it is necessary to prevent certain services (e.g. citizen-to-authority Emergency) sessions from monopolizing all available resources during events such as disaster, emergency, and DDoS attacks from impacting the availability of other priority services such as MPS and MCS.

Also, as 5G network is expected to operate in a heterogeneous environment with multiple access technologies, multiple types of UE, etc., it should support a harmonised QoS and policy framework that applies to multiple accesses.

Further, for QoS control in EPS only covers RAN and core network, but for 5G network E2E QoS (e.g. RAN, backhaul, core network, network to network interconnect) is needed to achieve the 5G user experience (e.g. ultra-low latency, ultra-high bandwidth).

6.7.2 Requirements

The 5G system shall allow flexible mechanisms to establish and enforce priority policies among the different services (e.g. MPS, Emergency, medical, Public Safety) and users.

NOTE 1: Priority between different services is subject to regional or national regulatory and operator policies.

The 5G system shall be able to provide the required QoS (e.g. reliability, end-to-end latency, and bandwidth) for a service and support prioritization of resources when necessary for that service.

The 5G system shall enable the network operator to define and statically configure a maximum resource assignment for a specific service that can be adjusted based on the network state (e.g. during congestion, disaster, emergency and DDoS events) subject to regional or national regulatory and operator policies.

The 5G system shall allow decoupling of the priority of a particular communication from the associated QoS characteristics such as end-to-end latency and reliability.

The 5G system shall be able to support a harmonised QoS and policy framework applicable to multiple accesses.

The 5G system shall be able to support E2E (e.g. UE to UE) QoS for a service.

NOTE 2: E2E QoS needs to consider QoS in the access networks, backhaul, core network, and network to network interconnect.

The 5G system shall be able to support QoS for applications in a Service Hosting Environment.

A 5G system with multiple access technologies shall be able to select the combination of access technologies to serve an UE on the basis of the targeted priority, pre-emption, QoS parameters and access technology availability.

The 5G system shall support a mechanism to determine suitable QoS parameters for traffic over a satellite backhaul, based e.g. on the latency and bandwidth of the specific backhaul .

NOTE 3: The case where a backhaul connection has dynamically changed latency and/or bandwidth needs to be considered.