ngmn

the engine of broadband wireless innovation

5G End-to-End Architecture Framework v2.0

5G End-to-End Architecture Framework

by NGMN Alliance

Version:	**V2.0.0**
Date:	**12-Feb-2018**
Document Type:	**Final Deliverable (approved)**
Confidentiality Class:	**P - Public**
Authorised Recipients: (for CR documents only)	

Project:	**P1-Requirements and Architecture**
Editor / Submitter:	**Adrian Neal**
Contributors:	**Adrian Neal (Vodafone), Sebastian Thalanany (U.S. Cellular), Steve Tsangkwong U (Orange), Richard Mackenzie (British Telecom), Peter Hedman (Ericsson), Paul Muschamp (British Telecom), Ahmed Alsohaily (Univ. Toronto), Chen Wei (China Mobile) Dan Wang (China Mobile), Shahar Steiff (PCCW Global), Hans J. Einsiedler (Deutsche Telekom), Tayeb Benmeriem (Orange), Ines Riedel (Amdocs), Philipp Deibert (NGMN), Javan Erfanian (Bell Canada), Farooq Bari (AT&T), Jan Groenendijk (Ericsson), Srisakul Thakolsri (NTT DOCOMO), Min Zuo (China Mobile), Charles Hartmann (Orange), Paul Bradley(Gemalto), Steve Babbage (Vodafone), Marc Kneppers (Telus), Axel Nennker (Deutsche Telekom), Jovan Golic (Telecom Italia).**
Approved by / Date:	**NGMN Board, 26th February 2018**

The information contained in this document represents the current view held by NGMN Ltd. on the issues discussed as of the date of publication. This document is provided "as is" with no warranties whatsoever including any warranty of merchantability, non-infringement, or fitness for any particular purpose. All liability (including liability for infringement of any property rights) relating to the use of information in this document is disclaimed. No license, express or implied, to any intellectual property rights are granted herein. This document is distributed for informational purposes only and is subject to change without notice. Readers should not design products based on this document.

Abstract: Short introduction and purpose of document

This document delineates the requirements in terms of entities and functions that characterise the capabilities of an E2E (end-to-end) framework. Architectural perspectives and considerations associated with the service categories - eMBB, mIoT, URLLC - envisioned for 5G (Fifth Generation) underscore the delineation of the E2E framework requirements. These requirements are intended as guidance in the development of inter-operable and market enabling specifications for a 5G ecosystem. This version resolves the outstanding Security and Identity management issues from the first published version, adds new references associated with those parts, and makes no changes to the other sections.

Document History

Date	Version	Author	Changes
14/09/2016	V 0.1.0	Adrian Neal, Vodafone	First version
20/09.2016	V 0.2.0	Adrian Neal, Vodafone	Addition of Devices section
04/10/2016	V 0.3.0	Sebastian Thalanany, U.S. Cellular, Adrian Neal, Vodafone	Addition of text to Introduction and References sections
11/10/2016	V 0.3.1	Sebastian Thalanany, U.S Cellular, Adrian Neal, Vodafone, Richard Mackenzie, British Telecom.	Text in the Devices section.
06/12/2016	V 0.3.2	Adrian Neal, Vodafone, Sebastian Thalanany, U.S. Cellular.	Addition of text agreed as a liaison to 3GPP SA2, and in definitions and devices sections.
16/01/2017	V 0.4.0	Adrian Neal, Vodafone	Addition of agreed text to Sections 5 (Network Slicing) and 9 (Management and Orchestration).
27/01/2017	V 0.5.0	Adrian Neal, Vodafone	Addition of agreed text to Sections 4.2, 6.1, 6.1.5, 6.1.6, 6.2.2, 7.1, 8.1, 9.1, 11.1, 11.2, 13.1, and 15.
28/01/2017	V 0.5.1	Sebastian Thalanany, U.S. Cellular	Addition/revision of text in sections covering definitions, access, and core.
31/01/2017	V 0.5.2	Sebastian Thalanany, U.S. Cellular	Added document purpose, Sections 6.1 and 6.2 to support ultra-low latency, high reliability and availability usage scenarios.
06/02/2017	V 0.5.3	Sebastian Thalanany, U.S. Cellular, Ahmed Alsohaily, University of Toronto	Definitions for network slice blueprint, network slice instance, service instance, and network function have been provided. Revised NSP, SP, and added VNSP. Revised the Abstract and clarified abbreviations. Filled up Section 6.3, added reference [8] and

			provided minor edits in sections 3-6, 8, 10 and 12.
28/02/2017	V 0.5.4	Sebastian Thalanany, U.S. Cellular, Chen Wei, CMCC, Paul Muschamp, BT, Ahmed Alsohaily, University of Toronto	Updated text, based on comments from the call on February 23rd, 2017. Added Service-Based Architecture (Section 6.1), 5G RAN functional decomposition (Sections 6.3.1 and 11.1), Reference [12]].
13/03/2017	V 0.5.5	Sebastian Thalanany, U.S. Cellular	Qualified “minimised coupling” in Section 6.1. Added comment in Section 6.3.1, indicating an elevation of the requirements text, while not alluding to SDO options, or implementation.
28/03/2017	V 0.6.0	Adrian Neal, Vodafone.	Inclusion of agreed text from 23rd March call and contributions. Outstanding comments converted to Editor’s Notes.
06/04/2017	V 0.6.1	Adrian Neal, Vodafone.	Inclusion of agreed text to Sections 6.3.1.1, 6.4.4, 7.1-4, 8.1-5 and 11.3.
28/04/2017	V 0.6.2	Hans J. Einsiedler, DTAG, Shahar Steiff, PCCW Global, Sebastian Thalanany, U.S. Cellular, Adrian Neal, Vodafone	Editorial scrub Minor additions in 6.1, addition of 9.4 Federated Orchestration, added content in the General sections Section 6.2.1 Introduced the concept of “Microservices” as an enabling facet in the end-to-end framework, in Section 6.2.3. Addition of agreed content from 27th April conference call.
01/05/2017	V 0.6.3	Adrian Neal, Vodafone	Additional changes to Section 9.4
04/05/2017	V 0.6.4	Adrian Neal, Vodafone, Tayeb Benmeriem, Orange, Ines Riedel, Amdocs, Philipp Deibert, NGMN Office	Agreed changes to Sections 2, 6.1, 9.2, 9.3, 9.4, 9.5 and 15 from the 4th May conference call.
11/05/2017	V 0.6.5	Klaus Moschner, NGMN, Adrian Neal, Vodafone	Changes to Section on Federated Orchestration and renaming MANO to 5G E2E MANO. Addition of Bell Canada as contributor (from V0.6.4).
04/06/2017	V 0.6.6	Sebastian Thalanany, U.S. Cellular	Edited Section 9, and 9.1. Added E2E (End-to-End) definition Added text and diagrams – network slice management and orchestration, in Sections 9.2, 9.2.1, 9.2.2, and 9.2.3

07/06/2017	V 0.6.7	Sebastian Thalanany, U.S. Cellular	Updated E2E related definition. Included clarification text in Sections 9.2.1, 9.2.2, and 9.2.3
19/06/2017	V 0.6.8	Sebastian Thalanany, U.S. Cellular	Added and revised text in Section 5.1.
27/06/2017	V 0.6.9	Adrian Neal, Vodafone	Agreed updates to Sections 5.1, 9.1 from 22nd June 2017 conference call.
11/07/2017	V 0.7.0	Sebastian Thalanany, U.S. Cellular, Tayeb Benmeriem, Orange	Updated diagrams and text in Sections 4.2, 6.2.2, 9.2.4, 9.2.5, 9.2.6, 9.2.7, 9.2.8. and 9.2.9
13/07/2017	V 0.7.1	Sebastian Thalanany, U.S. Cellular	Merged updates from V 0.6.9r1.
21/07/2017	V 0.7.2	Adrian Neal, Vodafone	Inclusion of changes to Sections 1, 6.1.1, 6.1.2 and 6.2.2, and new Figure 2, from July 20th conference call.
27/07/2017	V 0.7.3	Adrian Neal, Vodafone	Editorial clean up agreed on 27th July conference call.
22/08/2017	V 0.7.6	Sebastian Thalanany, U.S. Cellular, NGMN Security Competence Team (SCT): especially, Min Zuo, China Mobile, Charles Hartmann, Orange, Axel Nennker, Deutsche Telekom, Paul Bradley, Gemalto, and Jovan Golic, Telecom Italia.	Merged agreements through interim revisions post August 3, August 10, 2017 reviews and calls. Incorporated NGMN Network Management and Orchestration team inputs on management and orchestration. Incorporated NGMN Security Competence Team input on the security and identity management sections resulting from several iterations, discussions and calls. Confining devices to customer-side, “unified customer database” removed. Editorial changes together with diagrams, references, and cross-reference updates.
01/09/2017	V0.7.9	Adrian Neal, Vodafone, Steve Babbage, Vodafone, SCT NGMN: especially, Charles Hartmann, Orange, Paul Bradley, Gemalto, and Jovan Golic, Telecom Italia, Alec Brusilovsky, Interdigital.	Agreements from 31st August call. Restructuring of Security and Identity management sections. Re-definition of devices as Endpoint/User devices.
03/09/2017	V0.7.10	Adrian Neal, Vodafone, Hans Einsiedler, Deutsche Telekom, Sebastian Thalanany, U.S.Cellular, Peter Hedman, Ericsson, Srisakul Thakolsri, NTT DOCOMO.	Update to Figure 1, Clarification of “VF” in Figure 2, further clarifications to Sections 5, 6, 8, 9, and 14 and new Abbreviations plus a new definition for X-Haul. All arising from the internal company

			review period and contributions to the 1st September conference call.
07/09/2017	V0.7.11	Adrian Neal, Vodafone, Peter Hedman, Ericsson, Srisakul Thakolsri, NTT DOCOMO, Farooq Bari, AT&T, Shahar Steiff, PCCW Global, Paul Muschamp, BT.	Agreements from the 7th September E2E Architecture Framework conference call. Revisions to Sections 6.4.3, 7.3, 11.2.3 and 16.
11/09/2017	V0.7.12	Shahar Steiff, PCCW Global, NGMN SCT: Jovan Golic, Telecom Italia, Steve Babbage, Vodafone, Charles Hartmann, Orange, Paul Bradley, Gemalto, Alec Brusilovsky, InterDigital.	Rewording in Section 11.2.3. Detailed revision of Sections 13 and 15 agreed on the 7th September SCT conference call.
15/09/2017	V0.8.0	Srisakul Thakolsri, NTT DOCOMO, Axel Nennker, Deutsche Telekom, Charles Hartmann, Orange, Adrian Neal, Vodafone.	Agreed changes to Section 9.2 from the 14th September conference call. Editorial clean-up including realignment of reference numbering and editorial comments by e-mail.
17/09/2017	V0.8.1	Adrian Neal, Vodafone.	Removal of two references not used in the text. Preparation for Board Approval.
09/02/2018	V1.1.0	SCT, NGMN: especially, Jovan Golic, Telecom Italia, Steve Babbage, Vodafone, Charles Hartmann, Orange, Paul Bradley, Gemalto, Marc Kneppers, Telus, Yogendra Shah, InterDigital	Sections 13.6 on security challenges and 15.1 on IDM open issues from V0.8.1 are resolved, updated and incorporated in Sections 13.1-13-5 and 15.1, respectively. Seven new references [24]-[31] are introduced.
12/02/2018	V2.0.0	Adrian Neal, Vodafone	Editorial cleanup and preparation for Board Approval after agreement from the joint E2E Arch/SCT conference call on 12th February.

1 INTRODUCTION

The purpose of this document is to provide a high-level framework of architecture principles and requirements that provide guidance and direction for NGMN partners and standards development organisations in the shaping of the 5G suite of interoperable capabilities, enablers, and services. It builds on the architectural concepts and proposals implied by the NGMN White Paper [1] and subsequent deliverables published by NGMN. It is anticipated that this document will have versions, beyond an initial version, to reflect additional forward-looking requirements and/or updates as needed.

The elements of functional virtualisation, shift of computing to the edges of the network, and leveraging of spectrum distribution and flexibility, are among the dominant themes that shape the 5G ecosystem [1]. Optimisation of operational and performance efficiencies, while creating and delivering an exceptional and customisable user experience is of paramount significance [2][3].

2 REFERENCES

[1] NGMN 5G White Paper v1.0, Feb. 2015.
[2] 3GPP TR 22.891: Study on New Services and Markets Technology Enablers, Release 14, v1.0.0, Sep. 2015.
[3] Recommendations for NGMN KPIs and Requirements for 5G, June 2016.
[4] NGMN Description of Network Slicing Concept v1.0.8, Sep. 2016.
[5] E.U. 5G-PPP project TRANSFORMER. http://5g-transformer.eu/
[6] NGMN 5G security recommendations Package #2: Network Slicing, Apr. 2016. https://www.ngmn.org/uploads/media/160429_NGMN_5G_Security_Network_Slicing_v1_0.pdf
[7] ITU-T, “The tactile internet,” ITU-T technology watch report, Aug. 2014.
[8] NGMN 5G security recommendations Package #3: Mobile Edge Computing / Low Latency / Consistent User Experience, Oct. 2016. https://www.ngmn.org/uploads/media/161028_NGMN-5G_Security_MEC_ConsistentUExp_v1.3_final.pdf
[9] NGMN Perspectives on Vertical Industries and Implications for 5G v2.0, Sep. 2016.
[10] Dmitry Namiot, Manfred Sneps-Sneppe, “On Microservices Architecture”, International Journal of Open Information Technologies ISSN: 2307-8162 vol. 2, no. 9, 2014, pp. 24-27.
[11] NGMN Project RAN Evolution: Multi-RAT Joint Radio Operation (MRJRO) v1.1, Mar. 2015.
[12] NGMN Project RAN Evolution: Further Study on Critical C-RAN Technologies v1.0, Mar. 2015.
[13] NGMN “5G Network and Service Management including Orchestration” v2.12.6. Mar. 2017.
[14] NGMN 5G security recommendations Package #1, May 2016.
[15] General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679), http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf
[16] Regulation on Privacy and Electronic Communications, https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-privacy-and-electronic-communications
[17] ETSI NFV-SEC 003: Security and Trust Guidance, Dec. 2014.
[18] ETSI NFV-SEC 009: Report on use cases and technical approaches for multi-layer host administration, Jan. 2017.
[19] ETSI NFV-SEC 012: System architecture specification for execution of sensitive NFV components, Jan. 2017.
[20] ETSI NFV-SEC-013: Security Management and Monitoring Specification, Feb. 2017.
[21] ETSI NFV-SEC 014: Security Specification for MANO Components and Reference points, May 2017.
[22] FIPS PUB 140-2: Security Requirements for Cryptographic Modules, available at http://csrc.nist.gov/publications/fips/fips140-2/fips1402.pdf
[23] FIDO alliance, https://fidoalliance.org/
[24] Z. Kotulski, T. Nowak, M. Sepczuk, M. Tunia, R. Artych, K. Bocianiak, T. Osko, and J.-P. Wary, “On end-to-end approach for slice isolation in 5G networks – Fundamental challenges,”

Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 783-792, Prague, 3-6 Sep., 2017.

[25] NIST.IR 8114: Report on Lightweight Cryptography, Mar. 2017 *https://www.nist.gov/programs-projects/lightweight-cryptography*

[26] GSMA, "Identity and Access Management Requirements, version 1.0," Nov. 13, 2017.

[27] ISO/IEC 29192: Lightweight Cryptography, 2012-.

[28] ISO/IEC 29167: Automatic Identification and Data Capture Techniques, 2014-.

[29] Ashutosh Dutta, "Security Challenges and Opportunities in SDN/NFV and 5G Networks," AT&T presentation, June 15, 2017 https://docbox.etsi.org/Workshop/2017/201706_SECURITYWEEK/06_5GSECURITY/S01/AT%26T_Dutta.pdf

[30] ETSI NFV-SEC 002: Cataloguing Security Features in Management Software, Aug. 2015.

[31] 3GPP TS 33.501: Security Architecture and Procedures for 5G System, Release 15, v0.6.0, Dec. 2017.

3 DEFINITIONS

E2E

End-to-End, which refers to communications between two endpoint devices or user equipment, across any arrangement of intervening administrative domains

Haptic Sense

Haptic sense is perception characterised by touch. This type of perception is associated with tactile sense (derived from the Latin: *Tangere* - to touch), and kinaesthetic sense (derived from the Greek: *Kinesis* – movement, and *Aesthesis* – perception), for example body movement.

Network Function (NF)

Processing functions in a network. This includes a variety of control plane, user plane, and service functions that span the layers of the protocol stack. (e.g. radio network functions, physical layer functions, Internet Protocol (IP) routing functions, applications etc.) [4].

Network Service Provider (NSP)

Entity that provides network access service, and owns related resources and functions (e.g. virtualised or physical) for providing network access. The resources and functions include spectrum, mobility and access management across heterogeneous and/or composite access networks, network management and orchestration, and network elements.

Network Slice Blueprint (NSB)

A complete description of the structure, configuration and the plans/work flows for how to instantiate and control the Network Slice Instance during its life cycle. A Network Slice Blueprint enables the instantiation of a Network Slice, which provides certain network characteristics (e.g. ultra-low latency, ultra-reliability, value-added services for enterprises, etc.). A Network Slice Blueprint refers to required physical and logical resources and/or to Sub-Network Blueprint(s) [4].

Network Slice Instance (NSI)

A set of run-time network functions, along with physical and logical resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the Service Instance(s). A network slice instance may be fully or partly, logically and/or physically, isolated from another network slice instance [4].

Proprioceptive Sense Proprioceptive sense is perception characterised by a combination of body position and movement. This type of perception pertains to stimuli that are sensed and generated within an organism.

Service Instance (SI) An instance is a run-time construct of an end-user service or a business service that is realised within or by a Network Slice [4].

Service Provider (SP) Entity that provides an application layer service. The entity may be a third-party, or an NSP.

Vestibular Sense Vestibular sense is perception characterised by balance. This type of perception pertains to sensing via a cavity or vestibule, typically associated with the inner ear, which affects the state of balance of the body.

Visual Sense Visual sense is ocular perception that characterises seeing. This type of perception pertains to sensing via the eye.

X-Haul A common flexible transport solution for future 5G access networks, which aims to integrate fronthaul and backhaul networks with all their wired and wireless technologies in a common packet based transport network under SDN-based (software defined networks) and NFV-enabled (network functions virtualization) common control.

4 HIGH LEVEL END-TO-END ARCHITECTURE

4.1 Background

4.2 High level architecture

NGMN envisions an architecture that leverages the structural separation of hardware and software, as well as the programmability offered by Software Defined Networks (SDNs) and Network Function Virtualisation (NFV). As such, the 5G architecture is a native SDN/ NFV architecture covering aspects ranging from endpoint/user equipment, (mobile/ fixed) infrastructure, network functions, value enabling capabilities and all the management functions to orchestrate the 5G system. Application Program Interfaces (APIs) are provided on the relevant reference points to support multiple use cases, value creation and business models.

The architecture includes layers above the network layer. It allows for federation between separately administered domains at the resource and service layers to realise end-to-end network and service slice instances where one or more service providers or network service providers are involved. This naturally implies federation of network management and service orchestration as well.

The overall arrangement of actors, in a virtualised framework for service orchestration, utilising network slicing as a foundational building block, in the context of one or multiple administrative domains, broadly referred to as domains is depicted in Fig. 1.

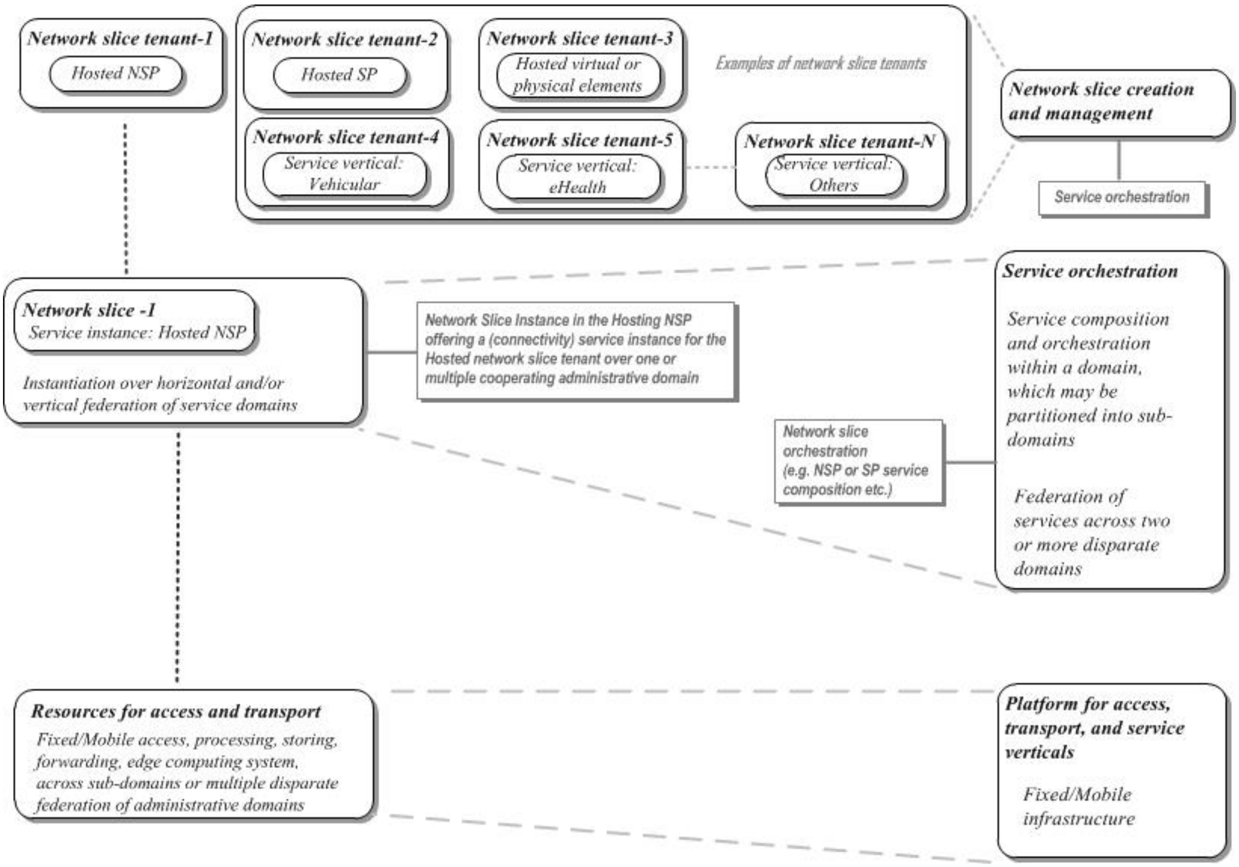
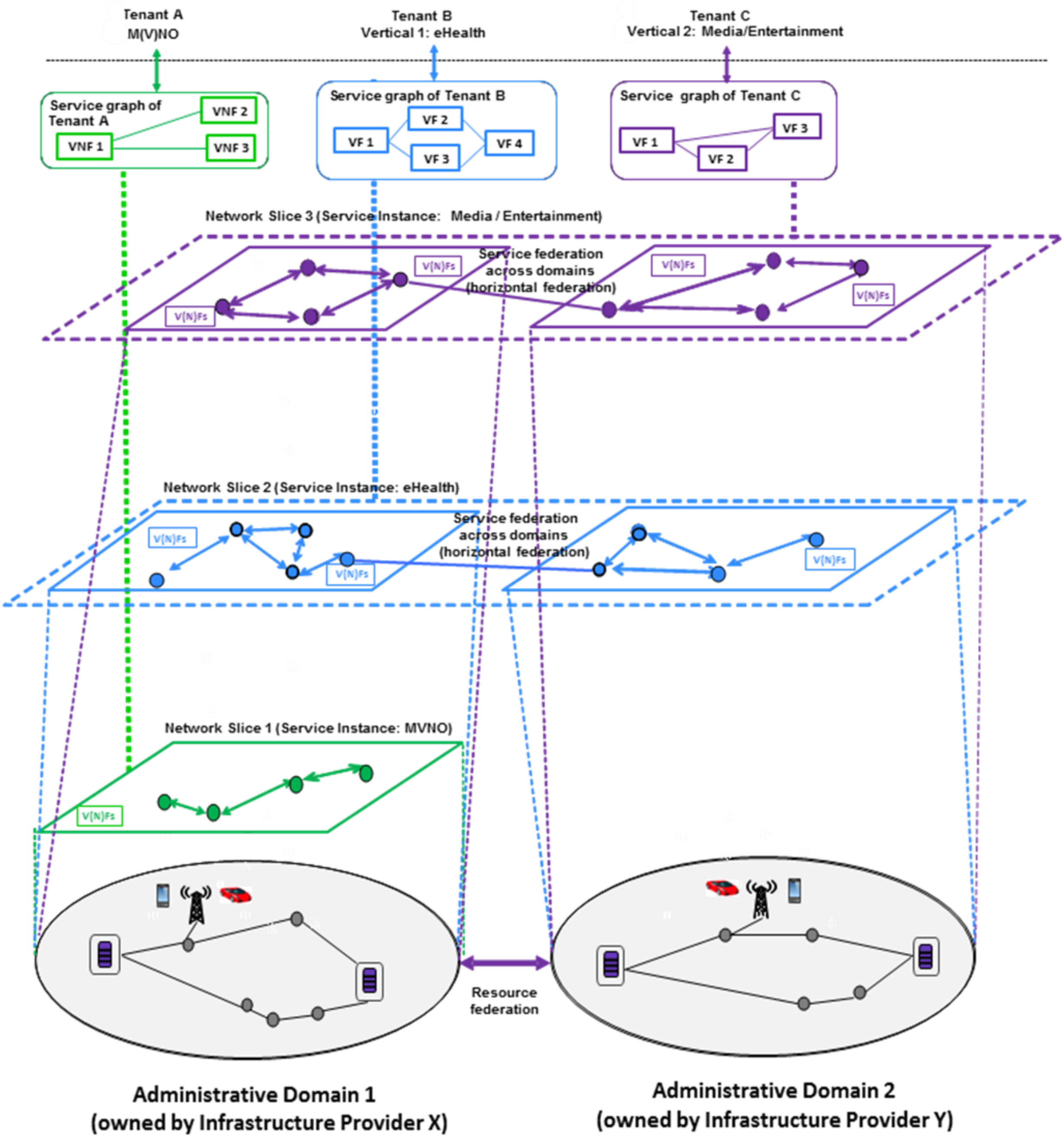

Network slice tenant-1
Hosted NSP
Network slice tenant-2
Hosted SP
Network slice tenant-3
Hosted virtual or physical elements
Examples of network slice tenants
Network slice tenant-4
Service vertical: Vehicular
Network slice tenant-5
Service vertical: eHealth
Network slice tenant-N
Service vertical: Others
Network slice creation and management
Service orchestration
Network slice -1
Service instance: Hosted NSP
Instantiation over horizontal and/or vertical federation of service domains
Network Slice Instance in the Hosting NSP offering a (connectivity) service instance for the Hosted network slice tenant over one or multiple cooperating administrative domain
Network slice orchestration (e.g. NSP or SP service composition etc.)
Service orchestration
Service composition and orchestration within a domain, which may be partitioned into sub-domains
Federation of services across two or more disparate domains
Resources for access and transport
Fixed/Mobile access, processing, storing, forwarding, edge computing system, across sub-domains or multiple disparate federation of administrative domains
Platform for access, transport, and service verticals
Fixed/Mobile infrastructure

Fig. 1 NGMN End-to-End 5G Framework vision.

Figure 2 presents an alternative view, depicting how federation of resources and services occurs between different administrative domains in order to provide the end-to-end service. Further detail is elaborated in the following sections of this document.



Tenant A
M(V)NO
Tenant B
Vertical 1: eHealth
Tenant C
Vertical 2: Media/Entertainment
Service graph of Tenant A
VNF 1
VNF 2
VNF 3
Service graph of Tenant B
VF 1
VF 2
VF 3
VF 4
Service graph of Tenant C
VF 1
VF 2
VF 3
Network Slice 3 (Service Instance: Media / Entertainment)
V(N)Fs
Service federation across domains (horizontal federation)
Network Slice 2 (Service Instance: eHealth)
Service federation across domains (horizontal federation)
Network Slice 1 (Service Instance: MVNO)
Resource federation
Administrative Domain 1
(owned by Infrastructure Provider X)
Administrative Domain 2
(owned by Infrastructure Provider Y)

Fig. 2 Resource and service federation across administrative domains

The European Union’s Horizon 2020 5G-PPP phase II project 5G-TRANSFORMER [5] has adopted the NGMN vision. The term VF refers to virtualised service functions at layers above the network layer.

5 NETWORK SLICING

5.1 General

The scope of a network slice is end to end. The 5G network shall be capable of slicing by service categories that consist of enhanced Mobile Broadband (eMBB, massive Internet of Things (mIoT), Ultra-Reliable Low-Latency Communication (URLLC), and other new arising categories.

The behaviour of a network slice, in terms of relevant resources, virtualised entities and functions, non-virtualised entities, at the user plane, control plane, and management/orchestration plane, is realised via the composition and instantiation of a network slice.

From an administrative domain perspective an end-to-end network slice may be within a domain, sub-domain, or across domains.

An end-to-end network slice is can involve more than one NSP or SP. A slice involving more than one NSP raises additional trust challenges and corresponding security requirements [6].

The notion of a domain or a sub-domain is within the jurisdiction of a single NSP or SP.

Multiple sub-domains are plausible within a single NSP or SP. Service categories may be sliced further. The extent to which a service category is sliced is established by the NSP.

A network slice may be composed of virtualised and/or non-virtualised entities.

More than one endpoint device or user equipment may connect to the same network slice (e.g. sensors and infotainment devices/user equipment for automotive).

The 5G system shall allow a common core network associated with one or more access networks to be part of a network slice (e.g. fixed and mobile access within the same network slice).

A Network Slice includes the following:

a) Control Plane functions associated with one or more User Plane functions (e.g. a reusable or common framework of control),
b) Service or service category specific Control Plane and User Plane function pairs (e.g. user specific multimedia application session).

Endpoint or user equipment can connect to a single network slice or to more than one slice. This raises additional security challenges, particularly relating to isolation between slices [6].

When endpoint or user equipment accesses multiple network slices simultaneously, a set of common control plane functions should be utilised by the multiple network slices and their associated resources.

The NSP (e.g. network operator or a virtual network operator, such as MVNO) uses a Network Slice Blueprint to create and manage a Network Slice Instance.

A network slice instance [4] may be;

Wholly statically defined, e.g., as in fixed-access business or residential service, or

Partially dynamic, e.g., as in roaming mobile endpoint/user equipment which may be connected to a statically-defined service chain, or

Fully constructed on demand;

Even when a network slice instance is statically defined, the necessary resources may be virtualised e.g., as transport tunnels over a layered infrastructure network, or as VNFs located somewhere in a cloud. The actual physical resources, together with their configuration, may thus vary over the course of time, including on-demand allocation or scaling.

A Network Slice Instance provides the network characteristics which are required by a Service Instance. Examples of a network slice instance include all the three categories of services – eMBB, mIoT, and URLLC – that span human-to-human, human-to-machine and machine-to-machine interfaces, which cover personal, industry, vehicular, social, health, city, and industry services and applications.

An example of a Sub-Network instance could be the IMS (IP Multimedia Subsystem), or a subset of network functions within an administrative domain realising parts of the network slice instance.

The Service Instance Layer represents the services (end-user service or business services) which are to be supported.

Each service is represented by a Service Instance.

> NOTE: A Service Instance can either represent an SP or an NSP service. The SP or NSP services may be a 3rd party provided service.

An administrative domain refers to the scope of jurisdiction of a provider. A provider may obtain service capabilities from 3rd parties to enrich the services it provides to its end customers. A provider could also benefit from offering its spare capabilities or resources to a 3rd party.

A network service can be a single user connectivity service, NaaS (Network as a Service) such as a service instance, a network slice instance or a subnetwork slice instance offering for a business vertical that utilises forward-looking business models, or IaaS (Infra-structure as a Service).

The notion of a partnership between two providers is qualified in terms of the one which is hosting the service, and the one whose service is being hosted, as described in the NGMN Network Slicing Concept paper [4]. A formalised description of the roles that qualify the behaviour of a provider is as follows:

Provider-Hosted (P-Hosted): A service provider that provides services to e.g. end customers, which can negotiate with another provider (Provider-Hosting) based on a trust model, for the establishment of a hosted network slice instance or a hosted Sub-Network instance using functions and resources from the hosting domain.

> NOTE: The necessary resources, in the hosting domain, are allocated based on a configured SLA between P-Hosted and P-Hosting,

Provider-Hosting (P-Hosting): A service provider, which can negotiate with another provider (P-Hosted) based on a trust model, for providing the usage of functions and resources in the hosting domain towards the hosted domain.

> NOTE: The necessary resources, in the hosting domain, are allocated based on a configured SLA between P-Hosted and P-Hosting,

Different types of partnerships and sharing may be envisioned, with a variety of distinctions:

> Various levels of functional exposure are considered, as envisioned in Section 4.5.2 of the NGMN 5G whitepaper [1].
>
> 5G should provide an abstraction layer as an interface, where all types of in-networking functionality (control plane and data plane related) can be exposed to the application layer functions and/or service providers based on a service level agreement. The application/service provider will then be able to use a sub-set of the network capabilities in a flexible, configurable and programmable manner, and to use network resources depending on their service preference.
>
> Automated real-time negotiations, as well as manual acquisition imply different considerations.
>
> Static or dynamic configuration of a partnership

Partnerships or agreements may be based on one or more bilateral agreements for realising any set of multiple partnerships

A bilateral partnership or agreement is typically based on an SLA (Service Level Agreement) between two parties, where each of the two participating providers are enabled to provide the necessary resources for the realisation of a service instance or a network slice instance.

For scenarios where the services of a broker are leveraged, there would be a pair of bilateral SLAs in place, where the broker behaves as a trusted mediator for the realisation of a service instance or a network slice instance.

5.2 Network Slicing – single administrative domain

The NGMN Network Slicing Concept paper [4] contains the following provisions;

A Network Slice Instance may be utilised by multiple Service Instances provided by the network service provider. This offers economy of scale/less overhead.

NOTE: Whether there is a need to support utilisation of Network Slice Instances across Service Instances provided by different 3rd parties is for discussion in Standards Development Organisations (SDOs).

A Network Slice Instance may be composed of zero, one or more Sub-Network Instances, which may be utilised by other Network Slice Instances.

A Sub-Network Blueprint is used to create a Sub-Network Instance from a set of Network Functions, which run on the physical/logical resources.

A Network Slice Blueprint is used to instantiate a Network Slice Instance.

A “Network Slice Instance-X” may be derived from a composite “Network Slice Blueprint-PQ” that has constituent “Sub-Network Blueprint-P” and a “Sub-Network Blueprint-Q”.

The “Network Slice Blueprint-PQ” is inherited from the constituents ““Sub-Network Blueprint-P” and “Sub-Network Blueprint-Q”.

The “Network Slice Blueprint” may also be a simple composition of “Sub-Network Blueprints”, where there is no inheritance.

5.3 Network Slicing – multiple administrative domains

The NGMN Network Slicing Concept paper [4] contains the following provisions;

The network slice blueprint may include resources or service capabilities from other providers with which an SLA exists.

In general, there are two categories of scenarios where network services need to be provided across multiple service providers:

Roaming scenario: Individual users move from one provider (i.e. Home NSP), which is the P-Hosted domain to a network managed by another provider (i.e. Visited NSP), which is the P-Hosting domain. The services that a user requires while roaming needs to be specified in the SLA between the two providers. In this case the two providers, with an SLA, would be the P-Hosted domain (Home NSP), and the P-Hosting domain (Visited NSP), with the corresponding behaviours required to support the inbound roamers (e.g. using a service instance or network slice instance) by the P-Hosting domain. The notion of tenancy, where the appropriate functions are provided by a P-Hosted domain to a P-Hosting domain based on SLAs is allowed, to meet the service characteristics required by the P-Hosted domain.

Business verticals: When a business vertical service user's request cannot be met by the capabilities of a single SP, the SP may harness the necessary capabilities from another SP, based on an SLA between the two SPs. In this case the two SPs, with an SLA, would be the P-Hosted domain (Home SP), and the P-Hosting domain (Third-party SP), with the corresponding capabilities required by the P-Hosted domain obtained from the P-Hosting domain.

6 NETWORK LAYER

6.1 Architectural considerations

The 5G system shall support a service-based architecture design, which enables modularised network services. The service-based architecture and interfaces in the 5G control plane make the 5G network flexible, customisable, and independently deployable. NSPs can leverage service-based architecture design in 5G to manage and customise the network capabilities, e.g., by dynamically discovering, adding, and updating network services while preserving performances and backward compatibility (when required). The network services functionality should enable reusability and loose coupling across network services. The service-based protocols should be lightweight.

The 5G core and access networks are to be functionally decoupled to create an access technology agnostic architecture [1]. The objective of the 5G architectural framework is to provide the flexibility required to realise the 5G performance targets for different usage scenarios. For example, the reduction in network latency requires the placement of computing resources and storage at the edge of the network to enhance service experience. This implies flexible orchestration of compute and storage resources from centralised to edge/cloudlet infrastructure.

The tactile internet [5] is a significant area of forward-looking usage scenarios, under the category of ultra-reliable low-latency communication services. A notable requirement for enabling the tactile internet is to place the content and context bearing virtualised infrastructure at the edge of an access network. This direction provides innovative directions for new revenue sharing opportunities and collaborative business models across various flavours of SPs. Content, context, and mobility demands are vital ingredients required to suit the demands of reliability, availability, and low-latency. Low latency / edge computing mechanisms raise additional trust challenges and corresponding security requirements [8].

The dominant themes within the tactile internet exemplify requirements for reliability and availability that need to be met at the access network edge to suit a variety of services that engage a human-in-the loop, across human-to-machine and machine-to-machine interfaces. The multimedia services in the tactile internet landscape are required to enable haptic interactions with visual feedback that augment the audio-visual user experience. Other multimedia components that are required as relevant for enabling tactile-internet services over a human-machine interface include vestibular and proprioceptive sensory translations. The tactile internet services with these multimedia component augmentations are required to be rendered with imperceptible latency. Such interfaces include robotic and machine-learning systems, with usage scenarios that span industry automation, telepresence, integrative health, autonomous vehicles, education, smart grid, renewable energy, personalisation, entertainment, art, cultural enrichment, etc.

The end-to-end latency required [5] for a satisfactory experience of tactile internet services is in the region of one millisecond, which implies much lower latency contributions over the radio-link segment, and under mobile handover conditions.

Human perception is guided by the sensory apparatus, which provides a measure of the quality of experience of interactions with the surrounding environment. This enables a feedback loop for an adaptation to the environment or to modify the experience of the environment. In the context of the tactile internet a corresponding service is an example of the environment. For a consistent, intuitive, and natural service experience, the service must be adaptable to the response time of human sensory perception.

The requirements to enable these new types of services, with the simultaneous demands of ultra-low-latency, reliability, availability, and mobility present the most challenging class of services that must be supported by an architecture framework that is sufficiently generalised, flexible, scalable, adaptable, and extensible. These architectural considerations are required to be examined at the access and core network layers.

NGMN has published some relevant information relating to the needs of vertical industries in [9].

6.1.1 Consistent User Experience across access networks

The 5G system shall support a consistent service experience across 3GPP and non-3GPP access networks, including in scenarios involving hand off between heterogeneous access technologies. The services may however have to be adapted to access specific characteristics for example in terms of QoS. Operationally a consistent service experience can be facilitated for example by adoption of a common set of procedures and functions for AAA, QoS, Policy, session continuity etc.

User applications should always be connected to a RAT, combination of RATs and/or attachment points (or other user equipment acting as a relay in case of D2D), or combination of points of attachment to the network providing the best user experience without any user intervention, according to NSP/SP subscription and policy.

The 5G system shall be able to provide an Inter-RAT mobility service interruption time that does not degrade the user experience, including between 3GPP and non-3GPP access technologies. The interruption time (if any) as a KPI should be compliant with the expected and desired user experience for the targeted service. For instance, for voice service, the 5G system shall be able to deliver service continuity by ensuring Inter-RAT seamless mobility. Some non-3GPP access networks may provide substantially lower security assurance than 3GPP networks do. Additional assurance may therefore be needed when accessing some core network features or services [8].

6.1.2 Fixed-Mobile Convergence considerations

To enable flexible management and joint optimisation the 5G system requires harmonised fixed and mobile Network Management and Orchestration.

Harmonising different identity and authentication, QoS, policy and charging paradigms in cellular networks, (wireless) local access networks, and fixed networks is essential to enable the convergence of different access types, to facilitate the realisation of different business models and maintain a consistent user experience.

6.2 Potential enablers for meeting required Key Quality Indicators

6.2.1 General

There may be several different enablers which could assist in meeting the Key Quality Indicator (KQI) requirements. This section describes some which may need to be developed further by the relevant SDOs.

Capabilities dispersed throughout the end-to-end framework are required in order to meet diverse KQIs associated with the three main categories of services, namely, eMBB, massive IoT, and URLLC. For example, in the case of the eMBB category of services high data rates at appropriate levels of QoS are a critical requirement, with the associated KQI targets. In case of massive IoT or massive MTC (Machine Type Communications) massive scale, variable payloads of information, low-cost options, battery longevity, low-maintenance, resource constrained operation etc. are among the requirements, with related KPI measures. The combination of high-reliability and low-latency requirements are among the QoS profiles, associated with URLLC services, with the related set of KQIs.

Energy efficiency, virtualisation, transport efficiency, handover efficiency, self-organisation, and enhanced utilisation of resources in the core and radio access networks, are pivotal enabling capabilities.

6.2.2 Optimisations for Edge Computing and Fixed/Nomadic uses

Edge Computing and Nomadic / Fixed access (wireless or wired) are two of the key 5G usage scenarios. Optimisation of the 5G framework for such cases must therefore be considered. For example, in the case of edge computing, the changes in the geographical location of a point of attachment of endpoint/user equipment to an access network edge resulting from mobility would add more overhead with tunnelling in a functionally virtualised network, which would impair an ultra-low-latency dependent service experience. Hence a minimisation of tunnelling overhead or the avoidance of tunnelling overhead may be required to scale as needed, while satisfying the most stringent requirements associated with tactile internet services. Similarly, when the endpoint/user equipment is stationary in its lifetime such as in the case of fixed 5G wireless deployment, paging it and use of tunnelling for the sake of mobility support can be viewed as an unnecessary overhead and complexity.

6.2.3 Microservices

The notion of microservices for applications is a significant enabling concept in the end-to-end-framework. As the name implies, a microservice is a small autonomous service that has its own architecture, technology, and platform. This type of service lends itself to distributed realisation of applications. The service can be managed, deployed and scaled in an independent manner throughout its lifecycle. For example, a microservices enabler could be applied for a realisation of desired functionality or customisation associated with customer experience, data analytics etc. The service can also be constructed from other combinations of building-block applications [10].

The benefit of utilising the microservices concept is that it is an enabler to suit various types of business models and contexts, in a manner that complements other enabling facets of the end-to-end framework, such as virtualisation and edge computing. Further details and applicability of these concepts are for further study.

6.3 Access Networks

The 5G core network will support multiple access networks including both fixed and mobile. FMC (Fixed-Mobile Convergence) is considered important (covered by requirements in all the following sections). Additionally, the 5G system will support the use of non-3GPP access for off-loading and maintaining service continuity. The 5G network shall enable the placement of applications taking latency or relevance to a defined geographical area into account.

Multiple connectivity (e.g. through multiple access technologies, or different links associated with the same access technology), where available, shall be supported to optimise resource allocation and signalling.

6.3.1 Mobile Access Network

The 5G system will allow multiple Radio Access Technologies (RATs) to be deployed and enable the seamless introduction of new RATs along with the flexible management and joint optimisation of radio frequency resources [8]. The redundant duplication of RAN functions for different RATs should be avoided, potentially through the unification of common RAN functions for different RATs. The simultaneous utilisation of multiple RATs by system users should also be enabled.

The 5G system will also support flexible RAN structures including implementations based on Cloud principles and the placement of context awareness at the RAN edges (i.e. mobile edge computing). Both centralised and distributed implementation of RAN functions should be enabled to facilitate the realisation of various RAN implementations. In addition, support for various coverage layers and cell sizes spanning extreme long-distance covering macro cells to small cell radio access deployments is required.

6.3.1.1 RAN Decomposition

Functional decomposition of the radio network is required to meet the diverse information transport demands (high performance to low performance) and align them with the demands of next-generation service categories of eMBB, mIoT, URLLC and other new arising categories. To accommodate them a decomposition of the radio network protocol layer functions, across layer-1, layer-2, and layer-3 is required, in terms of the degree of centralisation or distribution [12].

This decomposition consists of placing more functions of the upper layers of the radio network protocol stack in distributed entities for high performance transport demands (e.g. high bandwidth, high-capacity, low-latency, low-jitter etc.) relative to a centralised entity. Scheduling optimisation at a centralised entity, for high performance transport across multiple distributed entities (e.g. base stations, remote radio heads etc.) for fast coordination is a critical requirement.

For relatively low performance transport, more of the upper layer of the radio network protocol stack is placed at a centralised entity to optimise the cost/performance trade-off, associated with the distributed entities.
This choice of functional split will determine the X-Haul capacity requirement and associated latency specifications and performance. This will impact the network architecture as it could determine the placement of nodes and distance between them. A higher layer split will be tolerant of a large latency from a RAN perspective, which may be excessive when low-latency services are considered. Therefore, bounds must be applied within the network architecture to enable a service provider to support low latency services.

A distributed RAN (D-RAN) with several functional splits will be supported by 5G. Figure 3 illustrates the configuration with co-located centralised unit (CU) and distributed unit (DU). All radio protocol layers are terminated within the cell site.

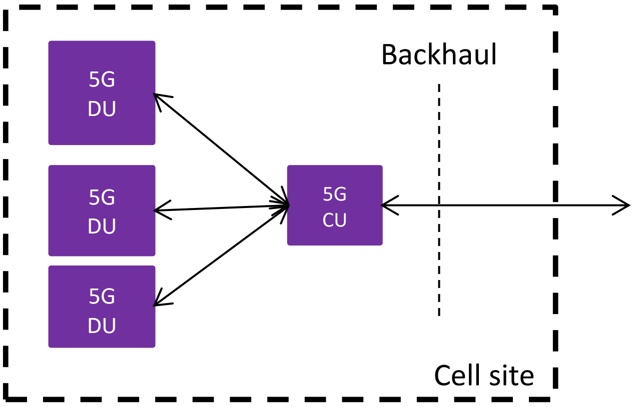


Fig. 3 D-RAN Configuration

The connection from the cell site towards the core network is traditional mobile backhaul which will be scaled and optimised to support 5G data rates and performance targets such as low-latency, low Packet Error Loss Rate (PELR), low and very deterministic Packet Delay Variation (PDV) etc. The D-RAN configuration does not constrain the ability of the local CU to support remote DU; in fact, the cell site could become a CU for other cells sub-tended as illustrated in Figure 4.

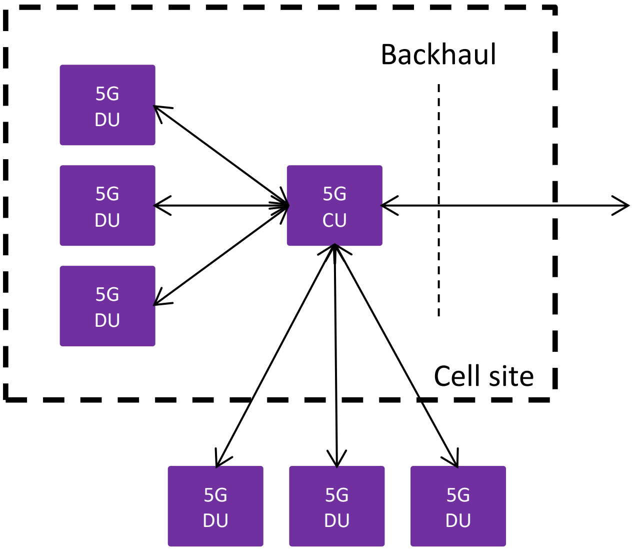


Fig. 4 D-RAN with sub-tended DU forming a local 5G C-RAN cluster with shared CU

A 5G C-RAN can be implemented with a higher layer split with the protocol stack with Packet Data Convergence Protocol (PDCP) being in the CU while the remainder of the stack is in the DU, as shown in Figure 5. This is one example; other splits will result in a different distribution of functionality between CU and DU.

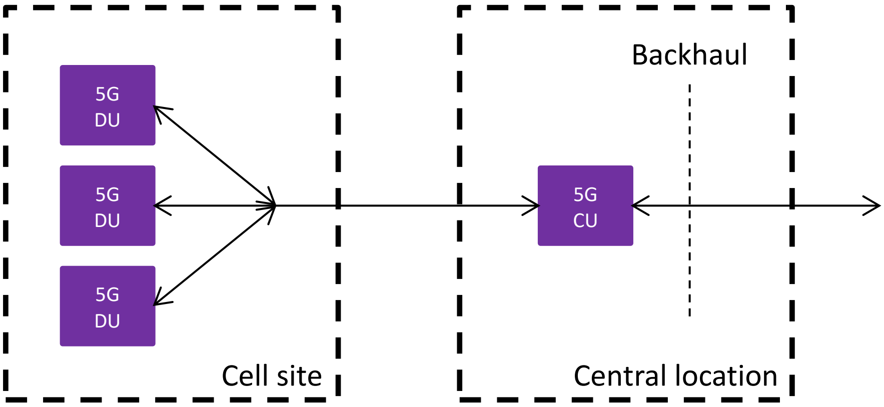


Fig. 5 C-RAN/D-RAN functional split

This configuration has similar X-Haul capacity requirements when compared with traditional backhaul. The latency and performance requirements of the RAN are not stringent and therefore consideration must be given to engineer the X-Haul link in accordance with service-based latency and performance targets.

6.3.2 Fixed Broadband Access Network

When applied to Fixed Broadband use cases the 5G system will provide provisions to improve user Quality of Experience (QoE) and maximise the efficiency of service delivery. Examples include Customer Premise Equipment (CPE) with higher capabilities than user equipment, reduced signalling to take advantage of the static placement of CPE and higher performing radio access configurations to exploit channel characteristics under static/outdoor CPE placement.

6.3.3 Wi-Fi Access Network

Among non-3GPP access technologies to be supported by 5G RAN is the 802.11 family, including current 802.11 releases (e. g. 802.11 ac and 802.11 ad) along with future releases (e. g. 802.11ax and 802.11ay). The 5G system shall provide provisions that ensure seamless access point integration, user access and mobility/flow management for Wi-Fi access technologies. This implies a need for automatic/SON-like solutions for fixed access management and orchestration.

6.3.4 Small Cells

The 5G system will enable the seamless integration of small cells under various deployment (such as planned NSP deployment and autonomous deployment) using wired or wireless backhaul. Autonomous deployment of small cells implies a need for automatic/SON-like solutions in small cell management and orchestration.

Small cells in the 5G system should be provided with effective interference cancellation means to enable their operation in the same frequency bands utilised by overlaying macro cells (i.e. co-channel interference) along with other bands not utilised by overlaying macro cells.

6.4 Core Network

The core network in the 5G system shall allow a user to access a network service, independent of the type of access technology.

6.4.1 General

The 5G core network will support multiple access networks including both fixed and mobile types of access networks.

The 5G system will provide termination points or points of attachment in the core, for both control plane and user plane information. These points are selected based on location, mobility, and service requirements. They may dynamically change during the lifetime of a service flow. To achieve a converged core network, common mechanisms of attachment should be supported for both 3GPP and non-3GPP access networks.

The 5G system will allow simultaneous multiple points of attachment to be selected for endpoint/user equipment, on a per-service flow basis.

The 5G system shall include a mechanism which provides network discovery and selection based on user experience, reliability and availability demands associated with the requested service.

6.4.2 Control and User Plane separation

Control and User Plane functions should be clearly separated with appropriate open interfaces defined among these types of functions.

6.4.3 NSP controlled Packet Data Networks

In 5G NSP controlled Packet Data Networks (PDNs) are connected to the User Plane Function (UPF) in 3GPP Core Networks via the N6 interface. Such PDNs may be virtualised and SDN controlled. In those scenarios the PDN functions (e.g. content filters, video optimisers, firewalls, DDoS protection etc.) shall be available as VNFs.

7 BUSINESS ENABLEMENT LAYER

7.1 General

The business enablement layer is a library of all functions required within a converged network in the form of modular architecture building blocks, including functions realised by software modules that can be retrieved from the library for use at a desired location, with an appropriate set of configuration parameters required for certain parts of the network, e.g., radio access. Figure 6 from [1] is illustrative of the context.

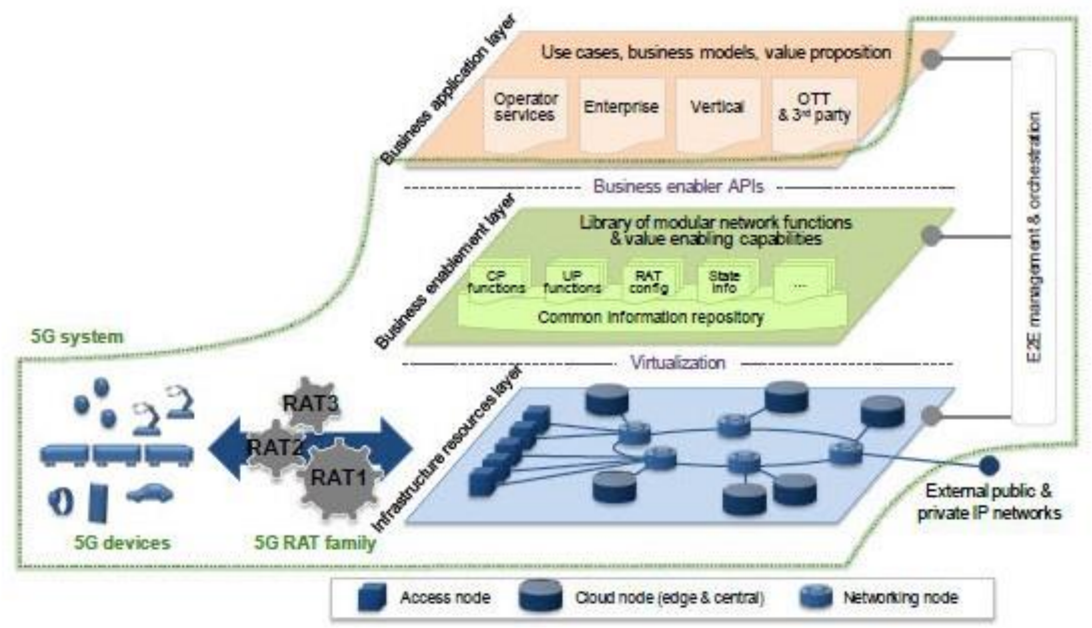

Business application layer
Use cases, business models, value proposition
Operator services
Enterprise
Vertical
OTT & 3rd party
Business enabler APIs
Business enablement layer
Library of modular network functions & value enabling capabilities
CP functions
UP functions
RAT config
State info
...
Common information repository
Virtualization
Infrastructure resources layer
5G system
RAT3
RAT2
RAT1
5G devices
5G RAT family
External public & private IP networks
E2E management & orchestration
Access node
Cloud node (edge & central)
Networking node

Fig. 6 5G system context

The functions and capabilities are called upon request by the orchestration entity, through relevant APIs. For certain functions multiple variants might exist, e.g., different implementations of the same functionality which have different performance or characteristics.
Business Enablement Layer functions shall be realised as virtualised network functions (VNFs) according to the principles specified by ETSI ISG NFV. Their lifecycle management and orchestration shall also be as described there, and their virtualisation requirements documented and managed according to the ISG NFV’s VNF Descriptor (VNFD).

3GPP VNFs shall be implemented as user plane/control plane specific 5G entities developed in alignment with Control/User Plane Separation.

7.2 Control Plane Functions

In an end-to-end multi-access network, the control plane functions are not limited to the 3GPP control plane. Control plane functions from fixed, WiFi and 5G mobile access and the converged 5G core network (including non-3GPP functions) are all in scope,

7.3 User Plane Functions

User plane functions include those from fixed, 5G mobile and WiFi access networks, the converged 5G core and NSP controlled PDNs connected to it, implemented as standalone user plane VNFs.

7.4 Configuration Data

Configuration data for each VNF is managed as per the procedures specified by ETSI ISG NFV. For 3GPP 5G functions it is managed as per the procedures jointly agreed between ISG NFV and 3GPP SA5. For non-3GPP functions equivalent processes analogous to the 3GPP versions are required.

8 BUSINESS APPLICATION LAYER

8.1 General

The business application layer contains the specific application packages and services of the NSP, enterprise, verticals or third parties that utilise the 5G network. In virtualised environments, it can be hosted in datacentres or on a Multi-access Edge Computing (MEC) host.

The interface to the management and orchestration system informs the Management and Orchestration (MANO) system of the required composition of dedicated network slices for an application, or the mapping of an application or service to existing network slices. The NGMN 5G White Paper [1] specifically left the detailed contents of this layer out of scope, as can be seen from Figure 1 of that document.

However, the interface to the end-to-end Management and Orchestration system is in scope. Management and orchestration for the application layer is required in a manner analogous to that for the Business Enablement Layer. The implication is that application and service layer software can be orchestrated and managed just like VNFs. Application and service software must therefore inform the Orchestration and Management system of its own infrastructure and runtime environment requirements just as a VNF would via the VNF Descriptor.
Accordingly, network slices can be created, orchestrated, and managed which contain all the physical and virtual network functions and application software required to deliver an end-to-end, multi-layer service.

It is desirable that applications conform to a standard, industry best practice, API format to ease their instantiation and to engage the widest possible community of application developers.

8.2 NSP Applications

NSP applications provide regular telecommunications services such as voice, messaging and internet access, as well as the NSPs own differentiating services which are offered to its own subscribers. The 5G system must include a mechanism whereby NSPs can rapidly instantiate, upgrade and remove new applications and new versions of existing applications, to trial new services and expedite upgrades or rollouts.

8.3 Enterprise Service Applications

NSPs offer service hosting to their enterprise customers. The 5G system must include mechanisms for enterprise service application packages, authenticated and authorised by the NSP, to be instantiated into the business application layer. From there they can form part of a bespoke end-to-end multi-layer enterprise service.

8.4 Vertical Service Applications

Some 5G use cases are realised by standalone private networks managed by the vertical industry itself rather than the NSP. A good example is factory automation. In such scenarios, the vertical can own and control its own application packages and business application layer. The 5G system must include mechanisms to enable this.

8.5 Authorised OTT and 3rd Party Service Applications

The 5G system shall include support for NSPs which offer service hosting for authorised 3rd party and OTT applications. The host can be a datacentre or MEC host. The 5G system must include mechanisms, by which the OTT player or 3rd party can request instantiation of, and management and usage reports from, their own applications.

9 END-TO-END MANAGEMENT AND ORCHESTRATION

9.1 General

NGMN's publication on 5G Network and Service Management including Orchestration [13] contains many specific requirements on the Management and Orchestration system. They cover such areas as Optimisation, Slice Management, Automation, and Self-organising functionality amongst many others.

The wide range of quality of service demands implicit in a 5G ecosystem demands a corresponding suite of enabling attributes that consist of interoperability, flexibility, extensibility, agility, and dynamism in the allocation of resources, within or across domains that leverage a virtualised environment.

9.2 Orchestration environment

The realisation of a wide range of quality of service demands over a virtualised environment requires the establishment of a network slice that utilises the appropriate resources necessary to support any given service realisation. A network slice may be established within an administrative domain (NSP or SP), or it may be established across multiple administrative domains (NSP or SP).

Service Instances [4] are managed and orchestrated at the Service Instance level.

All the constituent resources of a network slice instance are required to be visible to a 5G management and network orchestration sub-system, including non-virtualised resources (e.g. antenna elements, other elements at any layer of the protocol stack.) For a network slice that includes non-5G functions co-ordination between existing O&M systems and the 5G network management and orchestration system enables migration and coexistence strategies for a coherent and end-to-end management of a network slice.

The virtualisation of specific parts of functions, associated resources and elasticity should be managed at VNF/NFVI level by 5G management and orchestration implementations which are aligned with ETSI NFV MANO standards, to maximise interoperability. For each VNF application FCAPS management shall comply with the FCAPS management specifications published by the SDO which developed it, where possible.

Application and service layer software shall be orchestrated and managed in an analogous fashion to VNFs.

Application and service software shall inform the 5G management and orchestration system of its own infrastructure and runtime environment requirements just as a VNF would via the VNF Descriptor.

The 5G system shall include automatic/SON management and orchestration solutions for the deployment of non-3GPP technologies such as WiFi and their access points.

The 5G system shall include automatic/SON management and orchestration solutions for the autonomous deployment of small cells.

The 5G system shall be capable of managing and orchestrating network slice instances at the granularity of a network, Sub-Network or service. The interfaces for managing and orchestrating network slice instances and their constituent functions shall be open and standardised to enable interoperability of management and orchestration for the network functions and network slice instances.

10 POSSIBLE ORCHESTRATION ARCHITECTURE FLAVOURS.

The following architecture flavours are considered.

10.1 Vertical (Hierarchical) Orchestrator collaboration: layering view

Orchestration shall be multi-layer (vertical/hierarchical) in nature as it involves processes that start from the business level and inductively trigger lower level resource instantiations where synchronisation, delegation or escalation between orchestration layers may be needed. One possibility is that the actions of an orchestrator in one layer may also need to be synchronised with a higher-level orchestrator or for delegation / escalation purpose.

10.2 Federated Orchestration

When considering slices that are provisioned over multiple operators' networks or over multiple domains (Sub-Networks) within a single operator's network, an assumption of a single top-level orchestrator that has end to end visibility and control over all the domains and networks may not necessarily be true. This is more prominent across different operator/administrative domains, while in scenarios where the service is provisioned across technology domains operated by a single operator - hierarchical orchestration is more likely to be considered as an option. To construct such multiple domain service in the absence of a top-level orchestrator, the individual domain orchestrators must be federated in a manner that allows them to interface with each other horizontally for propagating slice policy and enforcing related rules. It is not necessarily involving hop-by-hop orchestrators along the orchestration path. This may imply some level of coordination / cooperation of autonomic decision-making aspect attached to orchestrators (Intent-based). In an environment where different domains may be operated using different controllers/orchestrators, the use of an industry-wide harmonised Information Model and industry wide standardised east-west-bound APIs is imperative.

10.3 Hybrid Federated and Hierarchical Orchestration.

Actual deployments may include a mix of federated and hierarchical orchestration where certain parts of the end-to-end service are orchestrated by a centralised orchestrator that controls the lower layers vertically, while such centralised orchestrators interface with their neighbour orchestrators horizontally in a federated manner.
Clearly the expectation is that regardless of the underlying method of orchestration, be it federated, hierarchical or a mix of both, the end user should receive ubiquitous experience, no matter how many operators may be involved in the delivery of service and the orchestration methods and approaches used.

11 NETWORK SLICE DEPLOYMENT MODELS

11.1 Categories of administrative domain configurations

The main categories of administrative domain configurations, for the establishment of a network slice, consist of the following:

a) Intra-domain
b) Inter-domain
c) Multi-domain

Each of these configurations can participate in the establishment of a network slice.

An intra-domain configuration refers to one or more sub administrative domains that are provisioned, within a single administrative domain to suit domain specific policies to handle different types of services. A network slice can be established to support a service, within a sub administrative domain. If a service requires the support of multiple sub administrative domains, then the network slice, required to support the service, is established through a cooperation of one or more designated orchestrators, based on the polices associated with the single administrative domain.

An inter-domain configuration refers to two different administrative domains that are required to cooperate to provide the necessary resources and functions to support any given service. The network slice required to support the service is established through a cooperation of the domain specific orchestrators, based on policies and agreements that are applicable across the two different participating administrative domains.

A multi-domain configuration refers to more than two different administrative domains that are required to cooperate to provide the necessary resources and functions to support any given service. The network slice required to support the service is established through a cooperation of the domain specific orchestrators, based on policies and agreements that are applicable across all the different participating administrative domains.

For each of these categories of administrative domain configurations, the participants in the establishment of a network slice would include a combination of SPs and NSPs, or just NSPs, depending on the nature of a given service realisation.

11.2 Network slice arrangements

A contextual view of the main categories of administrative domain configurations, for the allocation of the required resources for the establishment of a network slice, is shown in Fig. 7.

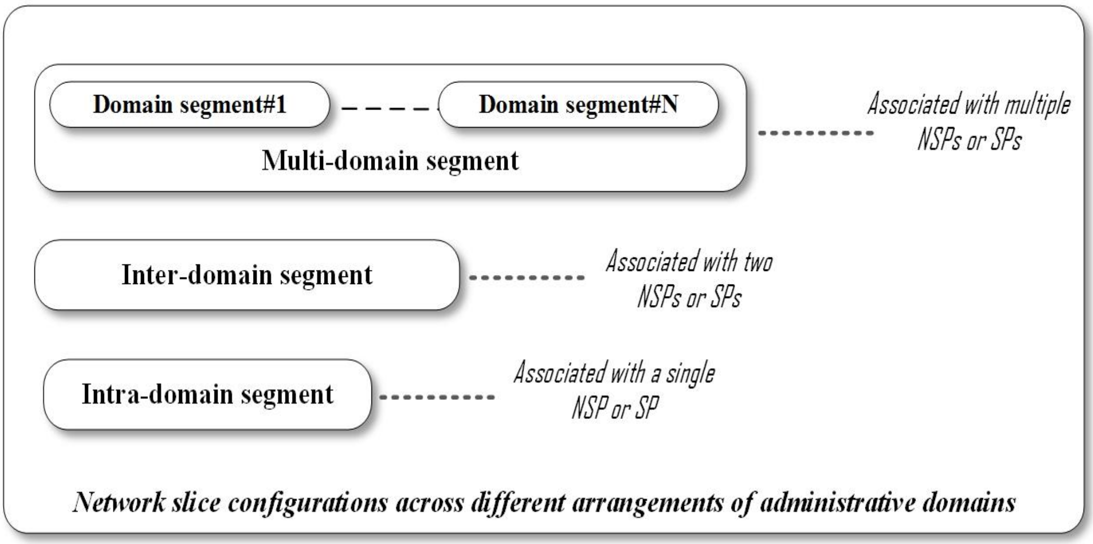
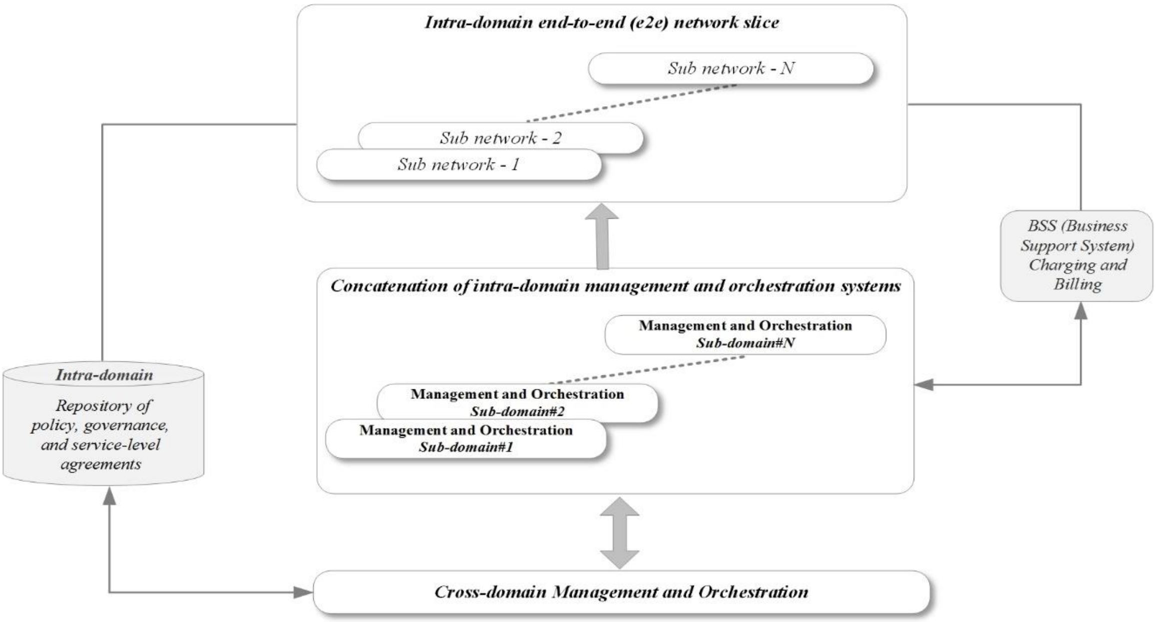

Domain segment#1
Domain segment#N
Multi-domain segment
Associated with multiple NSPs or SPs
Inter-domain segment
Associated with two NSPs or SPs
Intra-domain segment
Associated with a single NSP or SP
Network slice configurations across different arrangements of administrative domains

Fig. 7 Categories for network slice configurations

11.2.1 Intra-domain – E2E network slice

This scenario, shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**, depicts sub-domains, if present, within the same domain. Sub-domain management and orchestration systems are depicted as an arrangement of building-blocks from the inter-domain scenario, shown in Fig. 9.



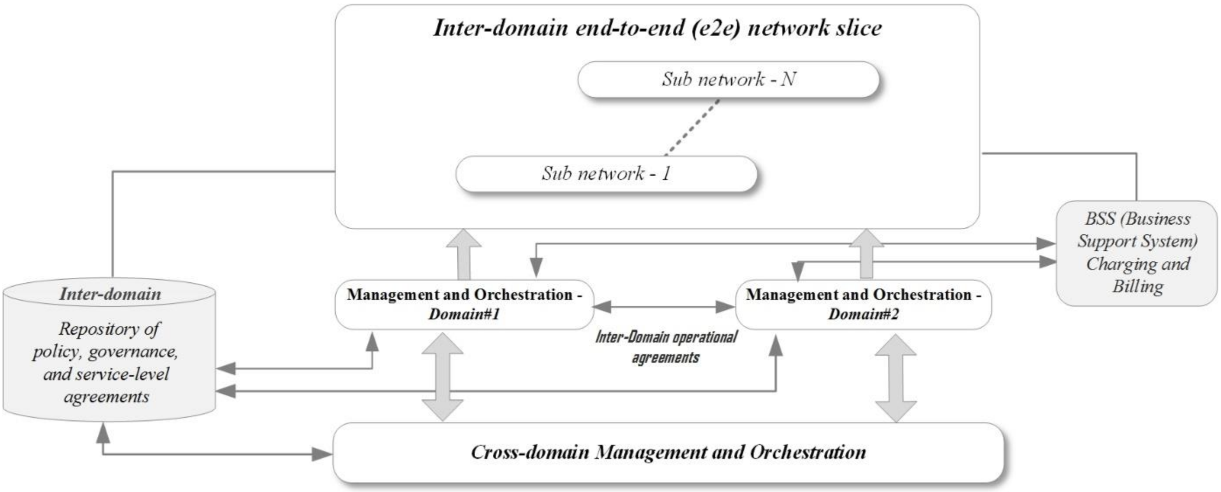
Intra-domain end-to-end (e2e) network slice
Sub network - N
Sub network - 2
Sub network - 1
BSS (Business Support System) Charging and Billing
Concatenation of intra-domain management and orchestration systems
Management and Orchestration Sub-domain#N
Management and Orchestration Sub-domain#2
Management and Orchestration Sub-domain#1
Intra-domain
Repository of policy, governance, and service-level agreements
Cross-domain Management and Orchestration

Fig. 8 Intra-domain E2E network slice

In **Fehler! Verweisquelle konnte nicht gefunden werden.**, any number of E2E network slices N, where N is a finite integer may be instantiated, through coordination across M different sub-domain management and orchestration systems, where M is a finite integer. The conceptual building-block, for management and orchestration across more than two sub-domains is derived and applied from the inter-domain scenario depicted in Fig. 9.

11.2.2 Inter-domain – E2E network slice

Fig. 9 shows the inter-domain scenario.



Inter-domain end-to-end (e2e) network slice
Sub network - N
Sub network - 1
BSS (Business Support System) Charging and Billing
Management and Orchestration - Domain#1
Management and Orchestration - Domain#2
Inter-Domain operational agreements
Inter-domain
Repository of policy, governance, and service-level agreements
Cross-domain Management and Orchestration

Fig. 9 Inter-domain E2E network slice

The scenario shown in Fig. 9, illustrates a conceptual view of the interactions among management and orchestration entities, where two different administrative domains are engaged in the establishment of a network slice.

In Fig. 9, any number of E2E network slices N, where N is a finite integer may be instantiated, through coordination between the management and orchestration systems in Domain #1 and Domain #2.

11.2.3 Multi-domain – E2E network slice

This scenario, in Fig. 10, is shown as an arrangement of building-blocks from the inter-domain scenario, shown in Fig. 9,

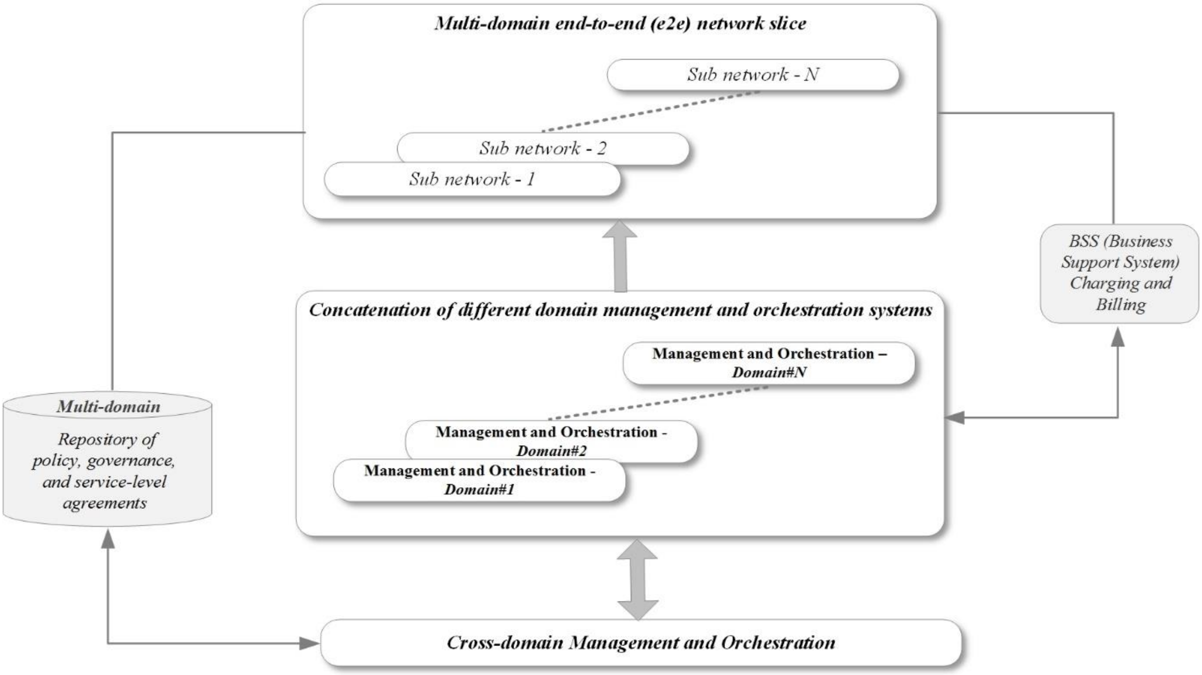

Multi-domain end-to-end (e2e) network slice
Sub network - N
Sub network - 2
Sub network - 1
BSS (Business Support System) Charging and Billing
Concatenation of different domain management and orchestration systems
Management and Orchestration – Domain#N
Management and Orchestration - Domain#2
Management and Orchestration - Domain#1
Multi-domain
Repository of policy, governance, and service-level agreements
Cross-domain Management and Orchestration

Fig. 10 Multi-domain E2E network slice

In Fig. 10, any number of E2E network slices N, where N is a finite integer, may be instantiated through coordination across M different inter domain management and orchestration systems, where M is a finite integer. The conceptual building-block, for management and orchestration across more than two domains is derived and applied from the inter-domain scenario depicted in Fig. 9.

The extra complexity of multi-domain (multi-lateral) network slice arrangements over inter-domain (bi-lateral) network slice arrangements introduces additional technical, operational and business challenges. Technical challenges include increased standardisation of the interfaces and models used, and how interconnection can be solved at the different layers. Operational challenges include process, information-model and data-model alignment. Business challenges include a need to align the minimum subset of capabilities that each administrative domain must have, having “standard” hosting/tenancy agreements (similarly to the roaming agreements of today), common settlement methodologies etc.

11.2.4 Federated Network Slicing Service Model

A service model that involves the participation of multiple domains, associated with two or more providers (NSPs or SPs), where a service is rendered for an end-user is depicted in Fig. 11. Each of the providers participates to allocate the appropriate functions and resources to compose a network slice to support "Service X", which is any given service for an end-user.

The use of a Federated Network Slicing (FNS) approach enables the creation of a network slice, where the network slice orchestrators in each of the participating domains cooperate to create a network slice for the creation of any service for an end-user. The FNS approach avoids the challenges associated with scalability, coordination, complexity, and the preservation of a consistent user experience, in the presence of end-user mobility, across disparate providers.

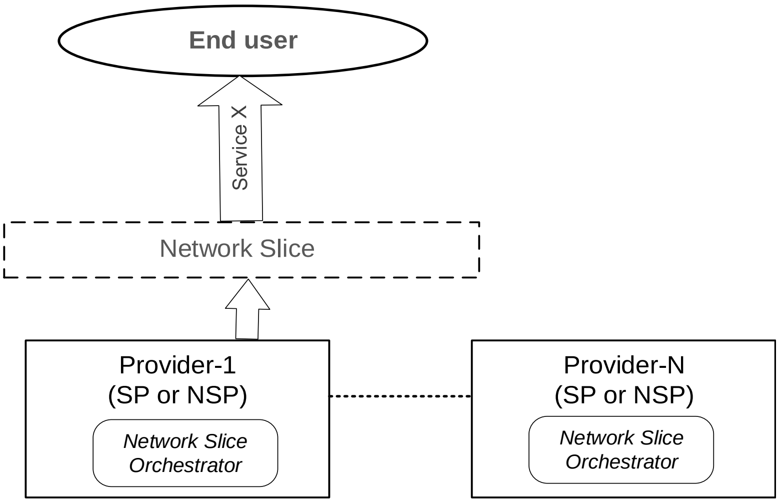


Fig. 11 Service model - Federated network slicing

FNS leverages the benefits of forward-looking capabilities, such as NFV/Virtualization.

As in scenarios such as home control/home routing or local breakout, in the FNS approach one provider (SP or NSP), is the designated service provider (P-Hosted) for an end-user. To complete the allocation of all the required functions and resources associated with a given service composition the engagement of other providers (SP or NSP) acting as partners (P-Hosting) may be needed, as required to suit the demands of a given service. For example, these other partners may provide capabilities such as connectivity or infrastructure as a service, using their NFV infrastructure etc.

In addition, a partner provider (P-Hosting) can provide connectivity services allowing the service provider (P-Hosted) to enable services via the partner (P-Hosting) for local access. The partner provider can in this case create a network slice used to provide services for the P-Hosted service provider. Hence a network slice within the partner provider network becomes a part of the service providers network slice (i.e. a network slice subnet).

To set this up the service provider orchestrator will, as part of setting up the network slice and its network slice subnet, send requests to the partner provider orchestrator, which will trigger the creation of the local network slice.

11.3 Inter-domain /Multi-domain / Intra-domain Slice Policy & Governance Repository

This repository stores the following items: Operator's Slice Policies, Governance, Objectives, Goals, negotiated SLAs/OLAs including compensations, commercial arrangements, responsibility demarcation among (Domains) and related roles and information to be exchanged at touch points.

Those items should be defined, negotiated, prescribed and updated in a flexible manner. Once appropriately combined and translated into desired rules, they will be propagated along the E2E Slice Orchestration path for execution and enforcement purposes.

11.4 The BSS (Charging & Billing) and its articulation with the operator's Slice Policy & Governance Repository

The execution phase includes usage tracking of services and its articulation with the BSS component for reporting, rating, charging per E2E Network Slice and Settlement per Subnetwork Slice provided (when applicable) and enforcement of the compensation rules in case of any SLA clause violation as prescribed in the SLA contract.

12 ENDPOINT/USER EQUIPMENT

12.1 Types

The types of endpoint/user equipment are characterised by a variety of attributes, within three broad categories of interfaces, namely, a) Human-Human (H-H), b) Human-Machine (H-M), and c) Machine-Machine (M-M). A few examples of user equipment or devices that belong to these categories are, smartphones (H-H), robots (H-M) or (M-M), drones (H-M) or (M-M), wearables (H-M), smart objects (M-M) etc. The attributes and capabilities, associated with them are diverse, such as, high-power, low-power, long battery life, low-cost, high performance, latency sensitive, high-reliability, precision sensitive. They are distinguished in terms of diverse media (synchronous, asynchronous, and isochronous) types, such as audio, visual, haptic, vestibular, data streams etc. They may be tethered to a network, either via a wired connection (e.g. Ethernet etc.), a wireless short range unlicensed connection (e.g. Wi-Fi, Bluetooth etc.), or wireless licensed connection (Cellular for instance).

12.2 Composite access

The availability of different types of RATs (Radio Access Technologies) for endpoint/user equipment to access a network or other endpoint/user equipment characterises composite access. For authenticated customer-side endpoint/user equipment session continuity between different types of access technologies, such as short range access technologies (e.g. Wi-Fi, Bluetooth etc.) and wide-area access technologies (e.g. different cellular schemes) is desirable. Composite access includes wired access, and different rates of mobility for wireless access. In such scenarios the availability of a diversity of access technologies, allows for optimisation of the utilisation of each access resource, as well as the selection of the most suitable access type for a given session.

Endpoint/user equipment may be connected to several access technologies (including new 5G RATs and LTE) at a given instant, potentially via carrier aggregation, or dual connectivity. The combination of access technologies may involve 3GPP access technologies and non-3GPP access technologies (licence exempt spectrum).

12.3 Heterogeneous access

Within a given type of wireless access technology the availability of different cell types with different characteristics, which inherently interoperate characterises heterogeneous access. For example radio network access elements, such as base stations with progressively larger to smaller coverage footprints are referred to as macro, pico and femto base stations respectively. A combination of these types of base stations offers the potential to optimise both coverage and capacity, by appropriately distributing pico and femto base stations within a larger macro base station coverage area.

Since the radio access technology is common across these different types of base stations common methods for configuration and operation such as 3GPP Licensed-Assisted Access (LAA) and optimisation of transmission power levels can be used to manage the allocation and utilisation efficiency of radio resources.

12.4 Cloud radio access

The cloud radio access model can be applied to composite and heterogeneous types of access. Resource offloading from endpoint/user equipment to the edge of the cloud radio access network enables diverse services over a variety of endpoint/user equipment types (e.g. H-H, H-M, and M-M). Other benefits may include energy conservation in the endpoint/user equipment and the ability to offer resource intensive services to the user even if computing or storage resources in the endpoint/user equipment are limited.

13 SECURITY

13.1 Network Layer Security

The 5G system shall support an access-agnostic subscription authentication framework capable of fulfilling the security requirements for both 3GPP and non-3GPP access.

The 5G system shall support subscription identities, for humans and machines/things, according to the international mobile subscription identity plan E.212 and possible future evolutions of this plan.

In 5G, the communication channels on control and user planes shall be integrity and confidentiality protected according to the security requirements building upon 4G and considering the specific role of the multi-access edge in MEC, which may require a combination of end-to-end and hop-by-hop techniques. The security requirements related to MEC [8] shall be respected.

With respect to low latency, the recommendations specified in [8] shall be met.

Trade-offs between ultra-low latency and security, especially with respect to confidentiality protection, may require usage of dedicated very fast cryptographic algorithms. In particular, light-weight, yet secure and trustworthy, cryptographic algorithms (e.g., ISO/IEC [27], [28], NIST [25]), may provide candidates, especially for constrained environments.

Network slicing, in particular E2E network slicing, is a central virtualisation technology in 5G which should help flexibly address the variety of use cases with different requirements as well as multi-vendor and multi-tenant network models over a shared hardware infrastructure. The security of network slicing shall meet the general requirements specified in [5]. The most important security-related property of network slicing is the isolation/separation of network slices, including the case where some network functions are allowed to be shared.

Fundamental challenges regarding the network slicing isolation are pointed out in [24]. There is a variety of isolation properties which can be satisfied at different isolation levels by using various technologies adapted to the desired isolation levels. The technologies include various software, hardware, and cryptographic mechanisms. In particular, the mechanisms may include orchestrator-managed containers, hypervisor-managed virtual machines, and virtual private networks. For each technology selected, it is important to specify concrete security requirements and the corresponding assurance levels.

If endpoint/user equipment is allowed to access one or multiple network slices then, before accessing any network slice, it shall be authenticated to access the 5G system via primary authentication described in TS 33.501 [31] and it should be authorised and/or authenticated for each network slice. A common authentication framework (e.g., EAP) should be used for this optional additional slice-specific authentication.

Massive IoT is likely to facilitate new variants of DoS and DDoS attacks on the network infrastructure and connected endpoint/user equipment. The corresponding security threats and recommendations given in [14] shall be respected.

In order to facilitate early detection and prevention of DoS/DDoS attacks initiated by malicious or infected massively distributed IoT devices, the IoT gateways, managing the devices and connecting them to the Internet, can be used for detecting anomalous or suspicious events. They should cooperate with classical network tools for DoS/DDoS detection and prevention based on analysing the aggregated traffic. New tools that will collect and analyse the RAN traffic may also be useful. This way the network infrastructure and IoT devices can be better protected in a timely manner.

Sensitive data (e.g., in e-health) generated by and transmitted among IoT/M-M devices should preferably be confidentiality, integrity, and anti-replay protected in E2E manner, possibly at the application layer.

In the 5G system, there are new security risks related to software and hardware implementation aspects in virtualised network infrastructure supporting virtualised network functions, network slicing, Service-Based Architecture (SBA), communication between applications, communication between virtualised network functions, and APIs. As a consequence, in order to address the security risks in 5G E2E architecture, in addition to the traditional network security approach based on protecting communication channels and protocols, a holistic approach, involving also computer security and cybersecurity aspects, is needed.

In the context of SBA, a study of the evolution of IPX, IPX provider services, and related security architecture should be performed.

An informative overview of the security challenges and opportunities in SDN/NFV and 5G networks is proposed in [29].

The 5G system should support privacy and security requirements compliant with international regulations, including general data protection and regulation [15], and with the proposal for a regulation on privacy and electronic communications [16].

The 5G system shall support relevant legal and regulatory obligations, including security aspects of Lawful Interception (LI). These legal and regulatory obligations may depend on country or regions.

The 5G system shall support protection of critical infrastructures, including the network infrastructure and 5G use cases such as Industry 4.0, e-health, Public Safety, Transport, Energy, Automotive, and massive IOT. The 5G system should consider compliance with the Directive NIS (Network and Information Security) (EU) 2016/1148 of the European Parliament and of the Council.

13.2 Business Enablement Layer Security

The Business Enablement Layer shall not compromise the 5G security architecture. Some relevant aspects of this layer, e.g., related to roles, responsibilities, and liabilities, are still to be defined.

The Business Enablement Layer may introduce new security requirements. This will apply especially to virtualised implementations (virtual appliances, hypervisors, OS, orchestrators, containers, etc.). See [8].

Since Business Enablement Layer functions shall be realised as virtualised network functions (VNFs) according to the principles specified by ETSI ISG NFV, the security considerations of these VNFs can be in alignment with the guidance published by ETSI ISG NFV-SEC in various aspects [14][18][19][20][30].

Whatever action/decision taken at this Business Enablement Layer, the NSP owning the network infrastructure should be in position to keep its sovereignty on its all network infrastructure. No third party could impose/force the NSP to import third-party software. Furthermore, the NSP, responsible for the network infrastructure, should be always able to remove / delete / isolate whatever VNF already put in production.

13.3 Business Application Layer Security

The Business Application Layer itself is out of scope according to [1].

The Business Application Layer shall not compromise the 5G security architecture.

Cross-application communication should be thoroughly and verifiably tested and strictly restricted to make sure the interface would not open-up new attack surface.

An NSP application shall be integrity protected throughout its lifetime, both in storage and desirably also in running.

Non-NSP applications, including enterprise service applications, vertical service applications, authorised OTT and third-party service applications, may introduce new security risks to the network [8]. The problem should be addressed by using a highly-secure virtualization platform with real-time security monitoring of Non-NSP applications (with anomaly detection and attack mitigation) or by running only trusted Non-NSP applications, previously tested according to the adopted quality assurance framework. In practice, security monitoring of Non-NSP applications can be simplified according to the risk analysis and vulnerability assessment conducted by the NSP.

In any case, the NSP should be in position to accept, deny, and remove any proposed application at any time and in any localization (NSP sovereignty on its own infrastructure). No third party or roaming agreement could force an NSP to implement a third-party application, even if the application has been validated by external framework/policy.

13.4 Management and Orchestration Security

Additional work and developments are needed to address the security concerns of some operators regarding the control of their network infrastructures depending on the technology / technical solutions used for virtualisation (e.g., based on containers, orchestrators, hypervisors, micro-services, DevOps). Specifically, opening interfaces/APIs at the orchestrator level, which are in control of the network infrastructure, creates new risk on the control of the operator network. Relevant threat analysis and security requirements are given in ETSI NFV-SEC 014 [21]. In any case, it is the responsibility of operators to make decisions on the virtualisation technologies to be used in their networks.

In general, due to its critical role in network virtualisation and network slicing, virtualisation management software should be highly secure and trustworthy. More specifically, this requirement relates to hypervisors for virtual machines and orchestrators for containers. In particular, virtualisation management software should be free of exploitable software vulnerabilities. Security and trustworthiness of virtualisation management software shall be supported by appropriate assurance levels.

The adopted solutions may depend on the obligations and requirements for critical infrastructures that emerge in certain countries or regions.

13.5 Endpoint/User equipment Security

Subscription data for the authentication of subscriber with the public 5G network shall be stored and processed in a tamper-resistant secure hardware component on the endpoint/user equipment. Solutions under development in 3GPP SA3 are consistent with this requirement. This requirement may possibly be extended under specified conditions to any persistently stored credential data needed for the authentication of subscriber with NSP/SP.

Secret or privacy-sensitive data for local authentication of a user to endpoint/user equipment (e.g., passwords or biometric templates) should be securely stored and processed on the endpoint/user equipment. For example, this security requirement can be implemented by using the solutions developed by the FIDO alliance [23].

Depending on the use cases, massive IoT may possibly require light-weight (yet secure and trustworthy) cryptographic algorithms that are adapted to power and bandwidth constraints, possibly the algorithms to be recommended by NIST [25] or proposed in ISO/IEC [27], [28].

In order to proactively protect the network infrastructure from new variants of DoS/DDoS and other attacks facilitated by massive IoT and to protect the IoT devices and the data generated by such devices, it is advantageous for IoT/M-M devices to run on simple, secure, and trustworthy CPUs and operating systems, with verifiable security levels. Preferably, these security levels should also reflect the sensitivity of data generated by such devices.

In order to proactively protect the network infrastructure from new variants of DoS/DDoS and other attacks facilitated by massive IoT and to protect the IoT devices and the data generated by such devices, authentication mechanisms for remote management of massively distributed IoT devices shall be secure and trustworthy.

A possible future threat is battery exhaustion attacks against devices – either indiscriminately at large scale or targeting individual devices whose disabling has particular value to the attacker. This is relevant primarily for IoT devices that run on batteries and cannot be easily or frequently recharged. If large scale battery exhaustion attacks become a significant problem, then network-based detection and prevention measures, including traffic analysis and management, may be needed. For individual devices whose availability is particularly important, defence against battery exhaustion attacks may be best implemented in the devices themselves.

The security of cryptographic functions also depends on their implementation, in hardware or software. To resist the so-called side-channel attacks (e.g., timing and power analysis attacks), it is necessary to apply appropriate countermeasures. For example, the requirements given in FIPS PUB 140-2 [22] could be used, with appropriate security level.

If endpoint/user equipment is allowed to access multiple network slices simultaneously, then isolation of the slices – and the ability to provide dynamic assurance of this isolation – should be provided by trustworthy software and/or hardware mechanisms based on an immutable root of trust, with appropriate security assurances, in order to prevent the leakage of (sensitive) data among the slices as well as the disruption of services in the slices. Note that this requirement refers to the mechanisms used to isolate execution and data storage on the endpoint/user equipment.

14 POLICY AND QUALITY OF SERVICE

14.1 General

The 5G system will support a common policy framework along with network policies that allow the endpoint/user equipment to choose the most suitable access network and an access agnostic quality of service mechanism. The 5G system shall support a common quality of service framework.

The common policy framework shall be access aware to enable conformance to service related QoS demands. In scenarios where more than one type of access (e.g. wireless, wired) is available, the choice of access hinges on the optimum (e.g. link conditions, efficiency, performance, policy etc.) suitability to satisfy QoS demands. The non-3GPP access solution could be a subset of the 3GPP access solution.

15 IDENTITY MANAGEMENT

15.1 General

The 5G system shall adopt a systematic approach to manage the following types of long-term identifiers:

- Endpoint/user equipment side:
    - Unique subscription identifier: e.g., IMSI/SUPI, etc.
    - Equipment identifier: e.g., IMEI
    - User (human or machine) identifier: e.g., username or pseudonym

  Unique subscription identifiers are owned and managed by operators. User identifiers are owned by users.

- NSP/SP side:
  - Network element identifier: e.g., a unique ID of a MME/AMF, etc.
  - Network device identifier: e.g., a femto ID associated with a digital certificate, etc.
  - Application identifier: e.g., a code-signing certificate, etc.
  - Service identifier: e.g., MSISDN, a NSP/SP certificate, etc.
  - Human identifier: including administrators, developers, etc.

A role of operators is to bind (e.g., during the registration) the unique subscription identifier to user-related, physically verifiable identifiers such as a username, where the corresponding secret key enables the remote authentication of the identifiers. Therefore, operators are in an advantageous position to perform the identity services also in more general usage scenarios providing new business opportunities.

Accordingly, the operators can provide user-centric identity services. More precisely, the operators can securely provision, manage, and share new identities, credentials, and authorization tokens linked to the subscription identity, provide remote authentication services based on them and control the privacy policies associated with them. The identities of humans may possibly include biometric identifiers. The related opportunities and requirements from GSMA [26] are under consideration in 3GPP. For IoT, to identify devices or things, it may be of interest to use Physical Unclonable Functions (PUFs) as biometric equivalents for integrated circuits. Operators and service providers must comply with privacy regulations on biometric and PUF data including protection against theft and reuse.

The 5G system shall support mutual authentication of an endpoint/user equipment-side identity and an NSP/SP-side identity or of two NSP/SP-side identities.

The 5G system shall adopt a systematic approach to manage short-term or temporary identifiers, for example, 5G equivalent of TMSIs, GUTIs, etc. Un-traceability: except for authorised entities, it should be hard to derive the long-term identifier from temporary identifier(s). Un-linkability: except for authorised entities, it should be hard to track the movement of the same endpoint/user equipment based on temporary identifier(s),

The use of strong user authentication, which is not based only on passwords or PINs is encouraged. For example, the solutions promoted by the FIDO alliance [23] may be useful for this purpose.

Long-term or permanent identifiers (e.g., IMSI/SUPI, IMEI, username) shall only be visible to authorised entities that need them for providing their function. They shall be stored securely and should not be transmitted in the clear. If they are transmitted encrypted, then the encryption shall be randomized for privacy reasons, to avoid linkability. The encryption/decryption key should be stored securely and the encryption/decryption operation should be executed in a secure environment.

16 LIST OF ABBREVIATIONS

3GPP	Third Generation Partnership Project
4G	Fourth Generation 3GPP system
AAA	Authentication, Authorisation and Accounting
API	Application Programming Interface
BSS	Business Support System
CPE	Customer Premises Equipment
CU	Centralised Unit
D2D	Device-To-Device
DDoS	Distributed Denial of Service
DoS	Denial of service
DU	Distributed Unit
EAP	Extensible Authentication Protocol

E2E	End-to-End
eMBB	Enhanced Mobile Broadband
ETSI	European Telecommunications Standards Institute
FCAPS	Fault, Configuration, Alarm, Performance and Security Management.
FMC	Fixed Mobile Convergence
GUTI	Globally Unique Temporary Identifier
GSM	Global System for Mobile Communications
GSMA	GSM Association
H-H	Human to Human
H-M	Human to Machine
IaaS	Infrastructure as a Service
IEEE	Institute of Electrical and Electronics Engineers
IMEI	International Mobile Equipment Identity
IMS	IP (Internet Protocol) Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IPX	Internetwork Packet Exchange
KPI	Key Performance Indicator
KQI	Key Quality Indicator
LTE	Long Term Evolution
M-M	Machine-to-Machine
MANO	Management and Orchestration
MEC	Multi-access Edge Computing
MIoT	Massive Internet of Things
MMTEL	Multimedia Telephony
N6	3GPP interface between the 5G core network and a Packet Data Network
NaaS	Network as a Service
NFV	Network Function Virtualisation
NFVI	NFV Infrastructure
NSP	Network Service Provider
OLA	Operations Level Agreement
OTT	Over-The-Top
PDN	Packet Data Network
QoE	Quality of Experience
QoS	Quality of Service
PDCP	Packet Data Convergence Protocol
PDV	Packet Delay Variation
PELR	Packet Error Loss Rate
PUF	Physical Unclonable Function
RAN	Radio Access Network
RAT	Radio Access Technology
SBA	Service-Based Architecture
SDO	Standards Developing Organization
SDN	Software Defined Networking
SLA	Service Level Agreement
SON	Self-Organising Network
SP	Service Provider
SUPI	Subscriber Permanent Identifier
TMSI	Temporary Mobile Subscriber Identity
URLLC	Ultra-Reliable Low Latency Communication
VNF	Virtualised Network Function
X-Haul	Flexible, heterogeneous access fronthaul and backhaul