ngmn
the engine of broadband wireless innovation

A Deliverable by the NGMN Alliance

NGMN 5G WHITE PAPER

next generation mobile networks

5G White Paper

By NGMN Alliance

Version:	1.0
Date:	17-February-2015
Document Type:	Final Deliverable (Approved)
Confidentiality Class:	P - Public

Project:	NGMN 5G Initiative
Editor / Submitter:	Rachid El Hattachi/ Javan Erfanian
Contributors:	5G Initiative Team
Approved by / Date:	NGMN Board/ 17th of February 2015

The NGMN White Paper has been endorsed by the following NGMN Board Members:

Tom Keathley
SVP Network & Product Planning

Bruce Rodin
VP Wireless Technology

Ameet Shah
Group Strategy Director

Li Zhengmao
EVP

Bruno Jacobfeuerborn
CTO

Jacob Groote
VP Mobile Operations

Seong-Mok Oh
Senior EVP & Head of Network Group

Seizo Onoe
EVP & CTO

Alain Maloberti
SVP

Tay Soo Meng
Group CTO

Alex Choi
CTO

Joachim Horn
Group CTIO

Sandro Dionisi
Director of Eng. & Telecom Italia Labs

Enrique Blanco Nadales
Group CTO

Günther Ottendorfer
CTO

Mats Svärdh
VP, Head of Group Networks and IT Infrastructure

Hugh Bradlow
Chief Scientist

Ibrahim Gedeon
CTO

Bülent Elönü
Chief Network Operations Officer

Yogesh Malik
Group CTO

Kevin Salvadori
Director Group Tech. Strategy & Operations

On behalf of the NGMN Alliance

Dr. Bruno Jacobfeuerborn
Chairman, NGMN Alliance

Dr. Peter Meissner
CEO, NGMN Alliance

NGMN 5G Initiative White Paper	
Leadership	Rachid El Hattachi (Deutsche Telekom) *Initiative Leader and Chairman of the Steering Committee* Javan Erfanian (Bell Mobility) *Co-leader of the 5G Initiative*
Editors	Javan Erfanian (Bell Mobility) *Chief editor* Brian Daly (AT&T) *Co-editor*
Work stream leaders	Armando Annunziato (Telecom Italia) Kevin Holley (Telefonica) Clark Chen (China Mobile) Eric Hardouin (Orange) Lou Feifei (China Mobile) Mikio Iwamura (NTT DoCoMo) Ralf Irmer (Vodafone) Stefan Apetrei (Orange) Chris Cheeseman (BT) Alex Tame (Vodafone)
Contributors	NGMN Members
PMO	Walter Grethe (Deutsche Telekom) Philipp Deibert (NGMN), contribution ‘way forward’ section Klaus Moschner (NGMN)
Steering Committee	Luigi Licciardi (Telecom Italia) Huang Yuhong (China Mobile) Takehiro Nakamura (NTT DoCoMo) Remi Thomas (Orange) Luke Ibbetson (Vodafone)

ACKNOWLEDGEMENTS

This NGMN 5G White Paper has been developed by NGMN's Members (i.e. operators) with contributions from NGMN's Sponsor (i.e. vendor) and Advisor (i.e. research) Partners. The White Paper has been delivered by a highly committed team and using a process of open contributions and continuous dialogue throughout the whole duration of the program.

We would like to express our appreciation to all who contributed directly or indirectly to the development of this White Paper.

The Board of NGMN has been of great support for the NGMN 5G global initiative by setting the overall direction, dedicating resources to the program and contributing to the overall content of the White Paper. The 5G team highly appreciates the level of commitment, synergy and drive by all Board members throughout the full duration of the program.

Our special thanks go to all members of the 5G initiative, and more specifically to the work stream leaders and co-leaders for driving their respective teams with passion and diligence. Our appreciation goes also to the Steering Committee members who contributed to the overall objectives of the program, by providing the necessary guidance to the team to deliver the White Paper in time and within the quality standard.

NGMN's Sponsor and Advisor Partners played a vital role in the delivery of the 5G White Paper. They were engaged and discussed numerous aspects of 5G at many sessions organised by NGMN's 5G initiative. The feedback given and the vision shared by those Partners were instrumental for developing the content of the White Paper. NGMN and the 5G team highly appreciate the commitment, effort and the valuable contributions from all Partners of NGMN.

Last but not least, the 5G team would like to express our sincere appreciation to the NGMN office and the PMO of the 5G initiative for the great support, contribution and dedication to the overall objectives of the 5G initiative as well as the overall preparation of NGMN's 5G activities at the Mobile World Congress and the NGMN Industry Conference and Exhibition.

Rachid El Hattachi – Deutsche Telekom
5G Initiative Leader &
Chairman of The Steering Committee

Javan Erfanian – Bell Canada
Co-Leader of the 5G initiative
& Chief Editor

TABLE OF CONTENT

LIST OF FIGURES AND TABLES

1. EXECUTIVE SUMMARY

The fifth generation of mobile technology (5G) is positioned to address the demands and business contexts of 2020 and beyond. It is expected to enable a fully mobile and connected society and to empower socio-economic transformations in countless ways many of which are unimagined today, including those for productivity, sustainability and well-being. The demands of a fully mobile and connected society are characterized by the tremendous growth in connectivity and density/volume of traffic, the required multi-layer densification in enabling this, and the broad range of use cases and business models expected.

Therefore, in 5G, there is a need to push the envelope of performance to provide, *where* needed, for example, much greater throughput, much lower latency, ultra-high reliability, much higher connectivity density, and higher mobility range. This enhanced performance is expected to be provided along with the capability to control a highly heterogeneous environment, and capability to, among others, ensure security and trust, identity, and privacy.

While extending the performance envelope of mobile networks, 5G should include by design embedded flexibility to optimize the network usage, while accommodating a wide range of use cases, business and partnership models. The 5G architecture should include modular network functions that could be deployed and scaled on demand, to accommodate various use cases in an agile and cost efficient manner.

In 5G, NGMN anticipates the need for new radio interface(s) driven by use of higher frequencies, specific use cases such as Internet of Things (IoT) or specific capabilities (e.g., lower latency), which goes beyond what 4G and its enhancements can support. However, 5G is not only about the development of a new radio interface. NGMN envisions 5G as an end-to-end system that includes all aspects of the network, with a design that achieves a high level of convergence and leverages today's access mechanisms (and their evolution), including fixed, and also any new ones in the future.

5G will operate in a highly heterogeneous environment characterized by the existence of multiple types of access technologies, multi-layer networks, multiple types of devices, multiple types of user interactions, etc. In such an environment, there is a fundamental need for 5G to achieve seamless and consistent user experience across time and space.

Business orientation and economic incentives with foundational shift in cost, energy and operational efficiency should make 5G feasible and sustainable. 5G should also enable value creation towards customers and partners through the definition and exposure of capabilities that enhance today's overall service delivery.

Enabling 5G use cases and business models require the allocation of additional spectrum for mobile broadband and needs to be supported by flexible spectrum management capabilities. In addition, an IPR eco-system needs to be developed to further enable innovation and unlock the potential associated with some of the use cases described in this paper.

NGMN and other stakeholders/partners will work together towards delivering globally and commercially available 5G solutions by 2020. This process will require a process of collaboration in the industry through existing standards development organizations (SDOs), or potentially new collaboration forms like open source.

2. INTRODUCTION

NGMN has had a central role in the definition of operator requirements, which has contributed significantly to the overall success of LTE. In the meantime, LTE has become a true global and mainstream mobile technology, and will continue to support the customer and market needs for many years to come. While supporting the development of LTE and its evolution, NGMN has developed the 5G requirements.

This is outlined by the operators, in close interaction with NGMN partners, in this White Paper. The NGMN White Paper serves as a guideline for 5G definition and design, and provides also insight into areas of further exploration by NGMN and other industry stakeholders.

This NGMN 5G White Paper begins with and builds on an outline of the 5G outlook. With the business context beyond 2020 distinctly different from today, the emergence of new use cases and business models are discussed. These are driven by the customers' and operators' needs and enabled by the maturity of existing and emergence of new key technologies. Based on this outlook and its attributes, the NGMN vision for 5G is formulated. This serves as an inspiration to develop the requirements and the related technology and architecture guidelines.

A detailed identification of the requirements follows the 5G outlook and the NGMN 5G vision. This is grouped along the six dimensions of user, system, device, service enhancement, network management and business requirements. To assess the capabilities expected to meet these requirements, first an analysis is provided in comparison with the current technology state of the art. A review of the technology trends provides insight on how the gap between the existing and the expected capabilities will be narrowed in coming years to a certain extent. Moreover, a perspective on potential technology building blocks that could further extend the capabilities and address the NGMN 5G requirements are outlined. At the heart of the discussion on technology and architecture, the reader finds a number of key design principles that are expected to be considered for the 5G architecture. A description of the 5G system architecture and components, with illustrative logical and physical realization examples and migration options serves as a guideline for further study and development within the NGMN program, standardization bodies, and across the ecosystem.

Enabling the complete 5G vision requires access to a wide amount and range of spectrum. The essential considerations for this are discussed, followed by an analysis of spectrum management aspects.

Guidelines towards a transparent and predictable IPR eco-system across the mobile industry are then identified to support the commercial implementation of 5G technologies and to ensure the enablement of a sustainable 5G eco-system.

Finally, this White Paper provides NGMN's view on the next steps. Support of standardisation, including open source, and opportunities to accelerate research, and technology and system innovations and solutions, are seen essential, as is the cooperation with all major stakeholders in the global eco-system around 5G.

3. 5G VISION

3.1 Business Context

Driven by technology developments and socio-economic transformations, the 5G business context is characterized by changes in customer, technology and operator contexts. It is expected that instant information will be just a touch away, and that everything will be connected.

Consumers

Significant recent technology advancement is represented by the advent of smartphones and tablets. While smartphones are expected to remain as the main personal device and further develop in terms of performance and capability, the number of personal devices will increase driven by such devices as wearables or sensors.

Supported by cloud technology, personal devices will extend their capabilities to various applications such as high quality (video) content production and sharing, payment, proof of identity, cloud gaming, mobile TV, and supporting smart life in general. They will have significant role in health, security, safety, and social life applications, as well as controlling home appliances, cars and other machines. To support such trends as multi-device and multi-access used by consumers, a comprehensive view of the future consumer's demands is essential.

Enterprises

Many of the trends in the consumer segment apply to future enterprises as well. The boundaries between personal and enterprise usage of devices will blur. Enterprises will look for solutions to address security and privacy challenges associated with this hybrid type of usage.

For enterprises, mobility will be one of the main drivers for increased productivity. In the next decades enterprises will increasingly make their specific applications available on mobile devices. The proliferation of cloud-based services will enable application portability across multiple devices and domains and will offer major opportunities for enterprises. At the same time this imposes challenges to enterprises that have to be managed properly (e.g., security, privacy, performance).

Verticals

The next wave of mobile communication is to mobilize and automate industries and industry processes. This is widely referred to as machine-type communication (MTC) and the IoT. Tens of billions of smart devices will use their embedded communication capabilities and integrated sensors to act on their local environment and use remote triggers based on intelligent logic. These devices differ in terms of requirements with respect to capabilities, power consumption and cost. IoT will also have a wide range of requirements on networking such as reliability, security, performance (latency, throughput), among others. The creation of new services for vertical industries (e.g. health, automotive, home, energy) will not be limited to connectivity but can require enablers from cloud computing, big data management, security, logistics and other network-enabled capabilities.

Partnerships

In many markets today, operators have already started to leverage partnerships with the so-called over-the-top (OTT) players to deliver packaged services to end users. OTT players will move to deliver more and more applications that require higher quality, lower latency, and other service enhancing capabilities (e.g., proximity, location, QoS, authentication) on demand and in a highly flexible and programmable way.

Infrastructure

Breakthrough technology advancements of the recent years (e.g., SDN, NFV, big data, All-IP) will change the way networks are being constructed and managed. These changes will enable the development of a highly flexible infrastructure that allows cost-efficient development of networks and associated services as well as increased pace of innovation. Operators will continue developing own services, but also expand their business reach through partnerships for both the infrastructure as well as the application development aspects.

Services

A global business model evolution of mobile operators' services will include the evolution of current services as well as the emergence of new ones. Currently the most common services provided by mobile operators include point-to-point personal communication and (best effort) data services such as Web services. These services will evolve to improve both in quality as well as in capability. Personal communication will include high quality IP multimedia and rich group communication as a baseline. Data services on the other hand, will be enabled by multiple integrated access technologies, will be ubiquitous, and will be characterized by performance consistency. Data traffic volume will be dominated by video and social media.

New services will emerge for growing and new market segments such as automated industries and smart user environments, public safety and mission critical services. Many other services will be developed by leveraging capabilities such as big data, proximity, geo-community services and many others.

3.2 5G Vision Characterisation

As outlined above, NGMN expects that the business context beyond 2020 will be notably different from today. The industry will see the emergence of new use cases and business models driven by the customers' and operators' needs. These will be enabled both by the maturity of existing and emergence of new key technologies. Therefore, NGMN has formulated the following vision for 5G that serves as an inspiration to develop the requirements and the related technology and architecture guidelines:

"5G is an end-to-end ecosystem to enable a fully mobile and connected society. It empowers value creation towards customers and partners, through existing and emerging use cases, delivered with consistent experience, and enabled by sustainable business models."

NGMN 5G Vision

The following sub-sections describe the key elements of NGMN's 5G Vision, namely:

- Use cases
- Business Models
- Value Creation

3.2.1 Use Cases

In addition to supporting the evolution of the established prominent mobile broadband use cases, 5G will support countless emerging use cases with a high variety of applications and variability of their performance attributes: From delay-sensitive video applications to ultra-low latency, from high speed entertainment applications in a vehicle to mobility on demand for connected objects, and from best effort applications to reliable and ultra-reliable ones such as health and safety. Furthermore, use cases will be delivered across a wide range of devices (e.g., smartphone, wearable, MTC) and across a fully heterogeneous environment. NGMN has developed twenty five use cases for 5G, as representative examples, that are grouped into eight use case families. The use cases and use case families serve as an input for stipulating requirements and defining the building blocks of the 5G architecture. The use cases are not meant to be exhaustive, but rather as a tool to ensure that the level of flexibility required in 5G is well captured. The following diagram shows the eight use case families with one example use case given for each family, and the description of these families and the use case examples are given below.

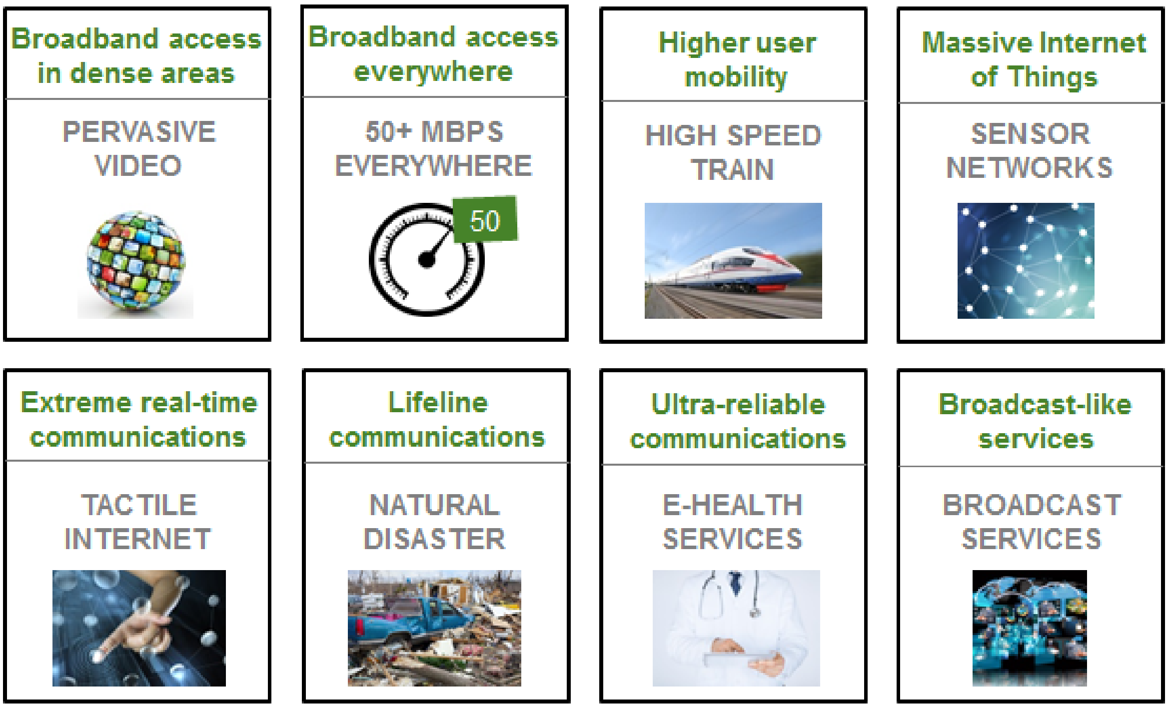


Figure 1: 5G use case families and related examples

Broadband Access in Dense Areas

This family highlights the broad range of growing and new use cases of the fully connected society. The focus is service availability in densely-populated areas (e.g., multi-storey buildings, dense urban city centres or events), where thousands of people per square kilometre (km^2) live and/or work. Communications are expected to be pervasive and part of everyday life. Augmented reality, multi-user interaction, three-dimensional (3D) services will be among the services which play an increasingly significant role in the 2020+ timeframe. Context recognition will be an essential aspect, at the network edge (i.e. close to the user), ensuring delivery of consistent and personalised services to the customers.

This family includes the following use cases:

i. Pervasive Video

Beyond 2020, person-to-person or person-to-group video communication with extremely high resolution will have a much wider usage with much more advanced and extreme capabilities. Customers will use video broadly in their everyday workflow. Examples include data delivery for optical head-mounted displays, collaboration in 3D cyber-real offices or operating rooms (with both physical and virtual presence) and customers' support by hologram services. An environment will emerge in which video is available to everyone, regardless of the physical location, the device being used, and the network connection. The number of concurrently active connections, combined with the performance required (data rate and the end-to-end latency) will present a challenging situation.

ii. Smart Office

In a future office, it is envisioned that most of the devices will be wirelessly connected. Users will interact through multiple and wirelessly connected devices. This suggests a scenario in which hundreds of users require ultra-high bandwidth for services that need high-speed execution of bandwidth-intensive applications, processing of a vast amount of data in a cloud, and instant communication by video. Ultra-high traffic volume, and for some applications latency, are the main challenges applicable for this use case.

iii. Operator Cloud Services

Cloud services provided by operators will become increasingly diversified, and further customized to each user, allowing operators to provide the user a full mobile "Smart life" experience. To support the future value added cloud services, there will be a need for higher QoE with user throughput consistency, fast and reliable networks, and seamless interworking across clouds, networks and devices.

iv. HD Video/Photo Sharing in Stadium/Open-Air Gathering

This use case is characterised by a high connection density and potentially temporary use (e.g., in a stadium, concert, or other events). Several hundred thousand users per km^2 may be served, possibly integrating physical and virtual information such as score, information on athletes or musicians, etc., during the event. People can watch high definition (HD) playback video, share live video or post HD photos to social networks. These applications will require a combination of ultra-high connection density, high date rate and low latency.

Broadband Access Everywhere

This family highlights the need to provide access to broadband service everywhere, including the more challenging situations in terms of coverage (from urban to suburban and rural areas). A consistent user experience with respect to throughput needs a minimum data rate guaranteed everywhere. Further development of digital inclusion of people living in scarcely populated areas and in developing countries requires the infrastructure deployment cost to be a key factor in services.

This family includes the following use cases:

i. 50+ Mbps Everywhere

The mobile and connected society will need broadband access to be available everywhere. Therefore, 50 Mbps should be understood as the minimum user data rate and not a single user's theoretical peak rate. Furthermore, it is emphasized that this user rate has to be delivered consistently across the coverage area (i.e. even at cell edges). The target value of 50 (or possibly 100) Mbps everywhere is meant to be indicative, depending upon the 5G technology evolution to support these figures economically.

ii. Ultra-low Cost Networks

Deployment and operation of mobile networks infrastructure as well as cost of terminals are not economically sustainable to cover scarcely populated and some very-low ARPU areas of the world. 5G is expected to be flexible enough to be deployed under ultra-low cost requirements to offer Internet access in these areas and enable new business and new opportunities in underserved areas of the world.

Higher User Mobility

Beyond 2020, there will be a growing demand for mobile services in vehicles, trains and even aircrafts. While some services are the natural evolution of the existing ones (navigation, entertainment, etc.), some others represent completely new scenarios such as broadband communication services on commercial aircrafts (e.g., by a hub on board). Vehicles will demand enhanced connectivity for in-vehicle entertainment, accessing the internet, enhanced navigation through instant and real-time information, autonomous driving, safety and vehicle diagnostics. The degree of mobility required (i.e. speed) will depend upon the specific use case.

This family includes the following use cases:

iii. High Speed Train

High speed train is used in various regions for inter-city transport and will further evolve beyond 2020; these high speed trains can reach speeds greater than 500 km/h. While travelling, passengers will use high quality mobile Internet for information, interaction, entertainment or work. Examples are watching a HD movie, gaming online, accessing company systems, interacting with social clouds, or having a video conference. Providing a satisfactory service to the passengers (e.g. up to 1000) at a speed of 500 km/h may be a great challenge. In addition, providing an acceptable end-to-end latency will become a challenge for office-like applications.

iv. Remote Computing

Beyond 2020, remote computing is used on the go and at high speeds (such as vehicles or public transport), in addition to those indicated for stationary or low-mobility scenarios (such as smart office). Moreover, automotive & transportation industry will rely on remote processing to ease vehicle maintenance and to offer novel services to customers with very short time-to-market. All this requires very low latencies with robust communication links together with availability close to 100%.

v. Moving Hot Spots

While moving vehicles or crowds (e.g., moving mass events such as walking/cycling demos or a long red-cycle of a traffic light) will generate capacity variation (from almost stationary to bursty), current radio planning determines hot spot areas, for optimization, assuming stationary hot spot. Therefore, non-stationary capacity demand will become a challenge in 2020+. 5G shall complement the stationary mode of planning of capacity, and incorporate non-stationary, dynamic and real-time provision of capacity.

vi. 3D (three dimensional) Connectivity: Aircrafts

Civil aviation will implement commercial connectivity services in 2020+, and the passenger services offered will comprise of similar applications to those available on the ground. Typical aircraft routes are up to 12 km in altitude, while other objects like helicopters will usually fly at much lower altitudes. Another example for 3D connectivity is support of sporting event live services where the user is moving physically in all 3 dimensions, e.g., balloonists, gliders, or skydivers.

Massive Internet of Things (IoT)

The vision of 2020 and beyond also includes a great deal of growing use cases with massive number of devices (e.g., sensors, actuators and cameras) with a wide range of characteristics and demands. This family will include both low-cost/long-range/low-power MTC as well as broadband MTC with some characteristics closer to human-type communication (HTC).

This family includes the following use cases:

vii. Smart Wearables (Clothes)

It is expected that the use of wearables consisting of multiple types of devices and sensors will become mainstream. For example, a number of ultra-light, low power, waterproof sensors will be integrated in people's clothing. These sensors can measure various environmental and health attributes like pressure, temperature, heart rate, blood pressure, body temperature, breathing rate and volume, skin moisture, etc. A key challenge for this use case is the overall management of the number of devices as well as the data and applications associated with these devices.

viii. Sensor Networks

Smart services will become pervasive in urban areas, and usage will also grow in suburban and rural areas. Among others, metering (e.g., gas, energy, and water), city or building lights management, environment (e.g., pollution, temperature, humidity, noise) monitoring, and vehicle traffic control represent prominent examples of services in a smart city. The aggregation of all these services leads to very high density of devices with very different characteristics expected to be combined in a common communication and interworking framework. Depending on the specific use cases, very low cost devices with very high battery life may be required.

ix. Mobile Video Surveillance

In the coming years, mobile video surveillance may evolve to be available on aircrafts, drones, cars, and safety and security personnel for monitoring houses/buildings, targeted areas, special events, etc. These applications will leverage automated analysis of the video footage, not requiring human support. While they will not present constraints on the battery life and often use medium/high-end devices, these applications require a highly reliable and secure network with the right performance and instant interaction with back-end and remote systems

Extreme Real-Time Communications

This family covers use cases which have a strong demand in terms of real-time interaction. These demands are use-case specific and, for instance, may require one or more attributes such as extremely high throughput, mobility, critical reliability, etc.. For example, the autonomous driving use case that requires ultra-reliable communication may also require immediate reaction (based on real-time interaction), to prevent road accidents. Others such as remote computing, with stringent latency requirement, may need robust communication links with high availability.

This family includes the following use cases:

x. Tactile Internet

Tactile interaction is referred to a system where humans will wirelessly control real and virtual objects. Tactile interaction typically requires a tactile control signal and audio and/or visual feedback. One application falling into this category is the use of software running in the cloud in a way that the user, interacting with environment, does not perceive any difference between local and remote content. Robotic control and interaction include countless scenarios such as those in manufacturing, remote medical care, and autonomous cars. The main challenge in tactile interaction is the real-time reaction that is expected to be within sub-millisecond.

Lifeline Communication

Public safety and emergency services that are provided today are continuously improving. In addition to new capabilities for authority-to-citizen and citizen-to-authority communication for alerting and support, these use cases will evolve to include emerging and new applications for authority-to-authority communication, emergency prediction and disaster relief. Furthermore, there will be an expectation that the mobile network acts as a lifeline, in all situations including times of a more general emergency. Therefore, the use cases require a very high level of availability in addition to the ability to support traffic surges.

xi. Natural Disaster

5G should be able to provide robust communications in case of natural disasters such as earthquakes, tsunamis, floods, hurricanes, etc. Several types of basic communications (e.g., voice, text messages) are needed by those in the disaster area. Survivors should also be able to signal their location/presence so that they can be found quickly. Efficient network and user terminal energy consumptions are critical in emergency cases. Several days of operation should be supported.

Ultra-reliable Communications

The vision of 2020 and beyond suggests not only significant growth in such areas as automotive, health and assisted living applications, but a new world in which the industries from manufacturing to agriculture rely on reliable MTC. Other applications may involve significant growth in remote operation and control that will require extreme low latency as well (e.g., enterprise services or critical infrastructure services such as Smart Grid). Many of these will have zero to low mobility.

This family includes the following use cases:

xii. Automated Traffic Control and Driving

In the coming years advanced safety applications will appear to mitigate the road accidents, to improve traffic efficiency, and to support the mobility of emergency vehicles (e.g., ambulances, fire trucks). These applications foresee not only a vehicle to vehicle or vehicle to infrastructure communication, but also communication with vulnerable road users such as pedestrians and cyclists. An application such as controlled fleet driving will require an ultra-low end-to-end latency for some warning signals, and higher data rates to share video information between cars and infrastructure. 5G should provide the high reliability, low latency, and high scalability required in this space.

xiii. Collaborative Robots: A Control Network for Robots

Automation will complement human workers, not only in jobs with repetitive tasks (e.g., production, transportation, logistics, office/administrative support) but also within the services industry. In order to enable these applications with completely diverse tasks in different environments, it will be essential to provide an underlying control network with very low latency and high reliability. For many robotics scenarios in manufacturing a round-trip reaction time of less than 1ms is anticipated.

xiv. eHealth: Extreme Life Critical

While mobile applications of remote health monitoring will continue growing beyond 2020, other applications such as remote treatment will emerge. Such applications will include several devices, like sensors, e.g., for electrocardiography (ECG), pulse, blood glucose, blood pressure, temperature. The monitoring applications, including the surveillance of patients remotely, will further grow in terms of availability and new applications. Depending on the patient’s device, treatment reactions may be required that are based on monitored data, and these should be immediate and (semi-)automatic. eHealth applications can be life critical and the system must be able to reserve/prioritise capacity for the related communications including out of coverage warnings. Identity, privacy, security and authentication management must be ensured for each device.

xv. Remote Object Manipulation: Remote Surgery

Remote surgery, available today using fixed networks, will be mobile in some scenarios such as in ambulances, for disaster-response, in remote areas, for the exploration of dangerous and hazardous areas, or during a leakage of radioactive material, etc. The technology necessary for providing the correct control and feedback for the surgeon entails very strict requirements in terms of latency, reliability and security.

xvi. 3D Connectivity: Drones

Future 5G services will require ubiquitous coverage, including both terrestrial and up-in-the-air locations. For example drones may be used for logistics such as autonomous delivery of packages on routes with no/low civil population. An example is delivery of medicine to the addressee, with drones automatically finding the way using a remote control system that exploits a 5G communication.

xvii. Public Safety

The public safety organisations will need enhanced and secure communications. This, for instance, will include real time video and ability to send high quality pictures. The main challenge is to ensure (ultra) reliable communication over the entire footprint of the emergency services including land, sea, air, in-building and some underground areas such as basements and subway systems. It will also require priority over other traffic (in networks shared with other users), ability for direct communication between devices, and high security.

Broadcast-like Services

While personalization of communication will lead to a reducing demand for legacy broadcast as deployed today, e.g. linear TV, the fully mobile and connected society will nonetheless need efficient distribution of information from one source to many destinations. These services may distribute content as done today (typically only downlink), but also provide a feedback channel (uplink) for interactive services or acknowledgement information. Both, real-time or non-real time services should be possible. Furthermore, such services are well suited to accommodate vertical industries' needs. These services are characterized by having a wide distribution which can be either geo-location focused or address-space focused (many end-users).

xviii. News and Information

Beyond 2020, receiving text/pictures, audio and video, everywhere and as soon as things happen (e.g., action or score in a football match) will be common. Customers in specific areas should simultaneously receive appropriate news and information regardless of the device they are using and their network connection.

xix. Local Broadcast-like Services

Local services will be active at a cell (compound) level with a reach of for example 1 to 20 km. Typical scenarios include stadium services, advertisements, voucher delivery, festivals, fairs, and congress/convention. Local emergency services can exploit such capabilities to search for missing people or in the prevention or response to crime (e.g. theft).

xx. Regional Broadcast-like Services

Broadcast-like services with a regional reach will be required, for example within 1 to 100 km. A typical scenario includes communication of traffic jam information. Regional emergency warnings can include disaster warnings. Unlike the legacy broadcast service, the feedback channel can be used to track delivery of the warning message to all or selected parties.

xxi. National Broadcast-like Services

National or even continental/world-reach services are interesting as a substitute or complementary to broadcast services for radio or television. Also vertical industries will benefit from national broadcast like services to upgrade/distribution of firmware. The automotive industry may leverage the acknowledgement broadcast capability to mitigate the need for recall campaigns. This requires software patches to be delivered in large scale, and successful updates to be confirmed and documented via the feedback channel.

The post-2020 outlook, shown throughout the use cases above, is extremely broad in terms of variety of applications and variability of their performance attributes. The use case families shown earlier represent both enriched service categories and also prospects for numerous new services. Note that some may have overlaps.

3.2.2 Business Models

On top of supporting the evolution of the current business models, 5G will expand to new ones to support different types of customers and partnerships. Operators will support vertical industries, and contribute to the mobilization of industries and industry processes. Partnerships will be established on multiple layers ranging from sharing the infrastructure, to exposing specific network capabilities as an end to end service, and integrating partners' services into the 5G system through a rich and software oriented capability set. There is a need for flexibility and embedded functionality to enable these. The following diagram shows examples of models that have to be supported by 5G.

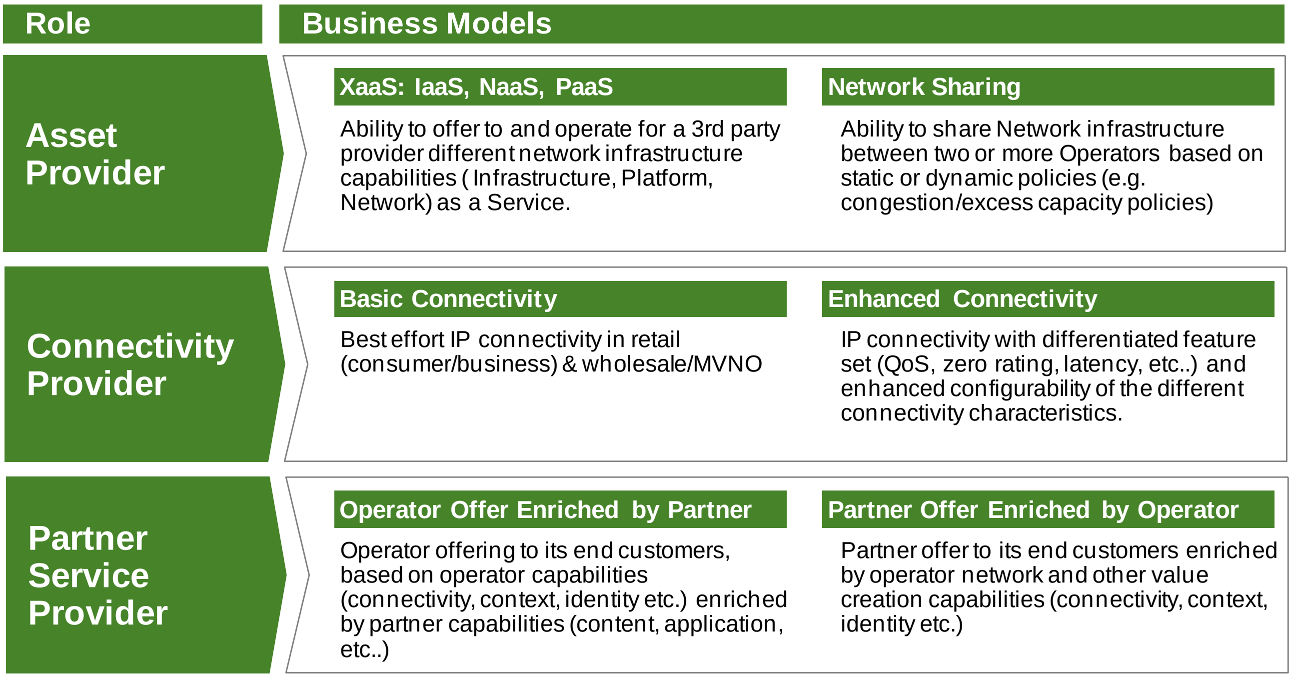


Figure 2: 5G Business models - Examples

Asset Provider

One of the operator's key assets is infrastructure. Infrastructure usually is used by an operator to deliver own services to the end-customer. However, especially in the wholesale business it is common that parts of the infrastructure – so-called assets - can be used by a third party provider. Assets can be different parts of a network infrastructure that are operated for or on behalf of third parties resulting in a service proposition. Accordingly, one can distinguish between Infrastructure as a Service (IaaS), Network as a Service (NaaS) or Platform as a Service (PaaS). These may be summarized as Anything as a Service (XaaS). Another dimension of asset provisioning is real-time network sharing that refers to an operator's ability to integrate 3rd party networks in the MNO network and vice versa, based on a dynamic and context dependent policies (e.g., congestion/excess capacity policies).

Connectivity Provider

Another role an operator can play in the future is one of a Connectivity provider. Basic connectivity involves best effort IP connectivity for retail and wholesale customers. While this model is basically a projection of existing business models into the future, enhanced connectivity models will be added where IP connectivity with QoS and differentiated feature sets (e.g. zero rating, latency, mobility) is possible. Furthermore, (self-) configuration options for the customer or the third party will enrich this proposition.

Partner Service Provider

Another role an operator can play in the future is one of a partner service provider, with two variants: The first variant directly addresses the end customers where the operator provides integrated service offerings based on operator capabilities (connectivity, context, identity etc.) enriched by partner (3rd party / OTT) content and specific applications. Integrated streaming solutions can be an example here but even services such as payments are possible.

The second variant empowers partners (3rd parties / OTTs) to directly make offers to the end customers enriched by the operator network or other value creation capabilities. Smart wearables with remote health monitoring are a good example. The customers buy clothes from a manufacturer and take benefit of the health monitoring feature offered by the 3rd party, enriched by the operator's set of network and value creation capabilities.

As a reflection of the above business context, the pricing models will also evolve and adapt to represent different types of services and customer profiles, for example,

- Evolved usage-based pricing, which reflects the throughput, latency, data consumption and device movement.
- Event based / real-time charging which may cover e.g. bandwidth consuming services.
- Tiered offers based on differentiated customer profiles and services.

Linked to the business context, the operators' capability to meet customers' demands, will depend on spectrum availability, roaming and assets sharing policies, and differentiated capabilities exposure. These impact operators' ability to develop new value propositions, and to provide quality service with consistent user experience throughout a wide range of scenarios. Therefore, it is evident that regulatory aspects will play a key role in 2020 and beyond

3.2.3 Value Creation

5G will bring multiple propositions to all customers and at the same time provides an enhanced and unique proposition tailored to each one of them. The definition of the customer is not limited to the consumers and the enterprises as in today's environment but also expand to include verticals and other partnerships. Common to all types of use cases and spanning all customer types, 5G will provide the following value proposition:

- **Available Anywhere-Anytime:**

  Delivering faster connectivity, communication and content anywhere, anytime without user perceived delay.

- **Delivered with consistent experience:**

  Services are delivered with a consistent experience across time, space, technology and devices used.

- **Accessible on multiple devices / interfaces:**

  User sessions are assumed to be portable from one device to another, in a transparent way to the user. Freedom to choose interfaces and forms of interaction (e.g., touch, speech, face and eye recognition).

- **Support multiple interaction types:**

  Multi-device interactions within smart user spaces and personal clouds with the user's ability to create, communicate, control, manage and share.

- **Supported transparently across technologies:**

  Full transparency and seamless connectivity for all customers regardless of the wireless or fixed accesses utilised.

- **Delivered in a personalised and contextual fashion:**

  Services are enhanced by contextual and personalized attributes to provide a personalised experience

- **Enabled by trusted & reliable communications:**

  Full trust, security and privacy supported.

- **Highly reliable and resilient network:**

  Mobile communication will be assumed to be always available as a lifeline, and serve as means for smart socio-economic well-being, smart services and processes, smart automated industries, and smart remote operations.

- **Responsive and real-time:**

  Extreme communication with stringent requirements, from fast downloads to real-time multimedia and pervasive video, with ultra-high resolution, for personal interaction and peer-to-peer or multi-party.

More specifically, **for consumers**, 5G will provide higher data rates and lower latencies required to support new and demanding applications. 5G as an engine of innovation will allow for faster development of new services delivered with consistent experience across time and space. Services and experience will be enhanced by contextual information leading to a very unique and personalized experience. On top of that, 5G should extend the battery life beyond today's norm.

For enterprises, 5G will provide differentiated capabilities to fulfil specific enterprise or enterprise application needs (security, privacy, reliability, latency, etc.). At the same time, through exposure of capabilities (e.g., location, analytics), enterprises can enrich and enhance processes and applications. Enterprise applications will enjoy the level of consistency of experience delivered by 5G.

For verticals, 5G will provide the required flexibility of functions and capabilities as it does for the enterprises. More specifically, 5G will provide the flexibility for verticals to operate their own applications in a profitable manner coming with a high degree of self service and at a cost level that allows sustainable business.

For 3rd party partners, 5G will foster innovation by flexible exposure of the network's value creation capabilities. This will enable partner-based propositions and allows for faster development and launch of these partner services at the benefit of all.

Operators' value creation propositions as outlined above will be enabled by capabilities that are flexibly integrated into the 5G system and easily exposed through APIs. This will be of significant benefit to all customers as it will allow for tailored and differentiated capability offering, enablement of new services, faster time to market and cost-efficient design.

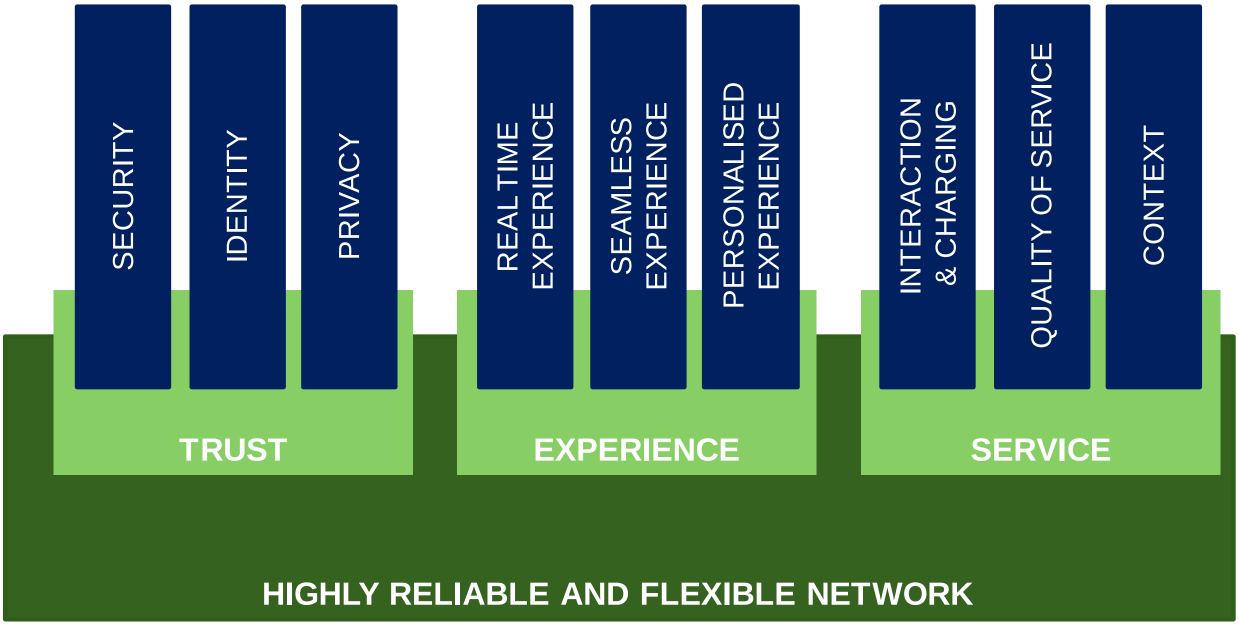


Figure 3: 5G Value creation capabilities

As depicted in Figure 3, on top of network connectivity, the value creation capabilities cover trust, experience and service related attributes. Trust includes capabilities such as security, identity management and privacy. Experience of services will be seamless and personalized across technologies, devices, time and location. From a service perspective, capabilities such as quality of service, context, and a responsive interaction and charging design will enable a differentiated service offering to customers and other service partnerships.

The following definitions describe the respective value creation capabilities in more detail. The value creation capabilities are expected to be embedded in the 5G design right from the start, and designed for exposure to enable a fast pace of innovation.

- **Security**

  The operator is the partner for state-of-the art data security, running systems that are hardened according to recognized security practices, to provide security levels for all communication, connectivity and (cloud) storage purposes.

- **Identity**

  The operator is the trusted partner for one (master) identity, providing for secured, hassle-free single-sign-on and user profile management to fit all communication and interaction demands.

- **Privacy**

  The operator is the partner to safeguard sensitive data, while ensuring their full handling transparency.

- **Real-time Experience**

  The operator enables perceived real-time connectivity to allow for instantaneous remote interaction among Things & People as if they are in close physical proximity.

- **Seamless Experience**

  The operator provides a seamless experience by managing and hiding the complexity involved in delivering services in a highly heterogeneous environment (multiple access technologies, multiple devices, roaming, etc.)

- **Personalized Experience**

  The operator is able to dynamically tailor delivered service experience based on customer context and a differentiated, customer configurable product portfolio.

- **Responsive Interaction and Charging**

  The operator is able to maintain a close relation with its customers throughout the lifecycle, by pro-actively triggering service or sales related transactions where and whenever relevant, and in real-time. This is enabled by a capability to identify events in real-time and apply the required business process in real-time (e.g. real-time charging)

- **Quality of Service**

  The operator is able to guarantee an agreed QoS, reliability and connectivity levels towards end customers and partners, over time and across the service coverage.

- **Context**

  The operator utilizes its contextual information asset to improve network operation and to enrich its service offering to end customers and partners.

4. REQUIREMENTS

The 5G requirements are derived out of NGMN's vision of the potential use cases and business models. Furthermore, NGMN believes that the requirements should also satisfy the value creation that operators intend to deliver to the different types of customers and partners. In summary, the NGMN vision leads to requirements that are grouped along the six dimensions shown below.

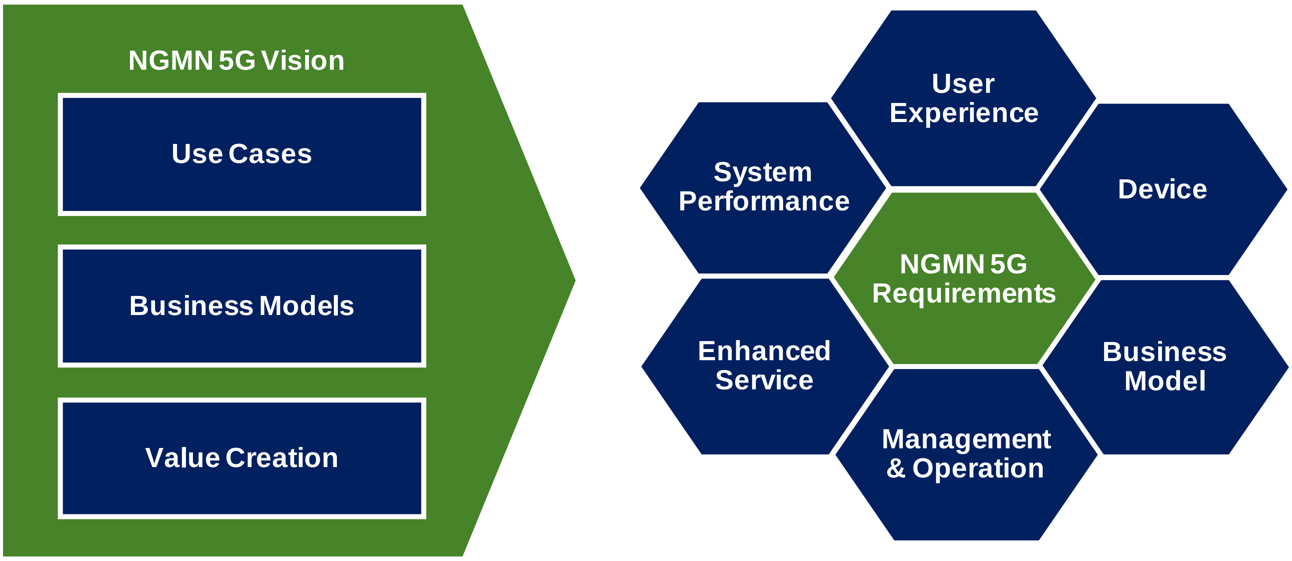


Figure 4: 5G requirements dimensions

The 5G use cases demand very diverse and sometimes extreme requirements. It is anticipated that a single solution to satisfy all the extreme requirements at the same time may lead to over-specification and high cost. Nevertheless, several use cases are anticipated to be active concurrently in the same operator network, thus requiring a high degree of flexibility and scalability of the 5G network. In order to reflect their use-case dependency, the requirements are specified according to the "Use Case Categories" defined in the figure below. For each use case category, one set of requirement values is given, which is representative of the extreme use cases(s) in the category. As a result, satisfying the requirements of a category leads to satisfying the requirements of all the use cases in this category.

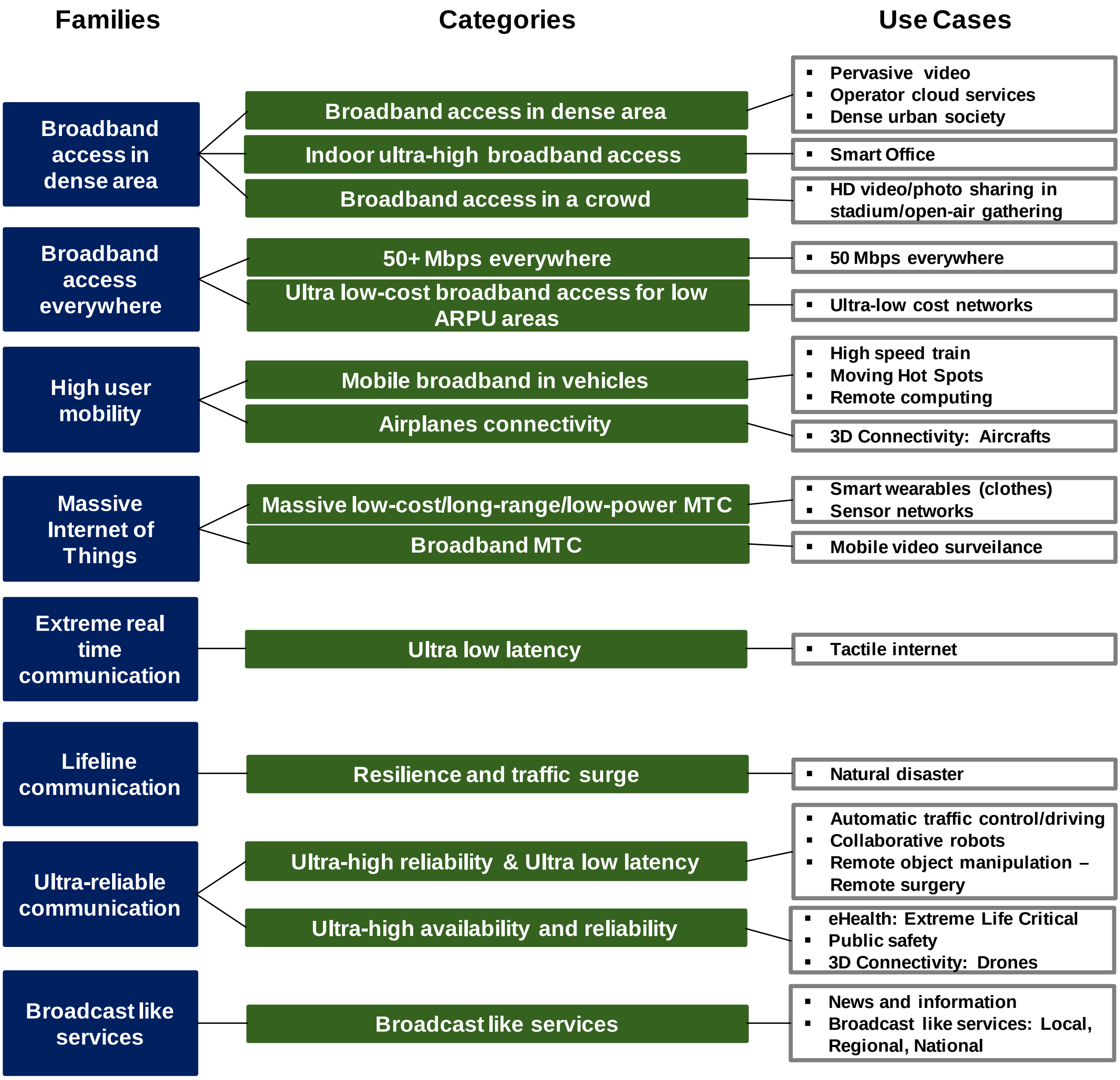


Figure 5: Use case categories definition

It is important to note that the NGMN requirements are based on the operator vision of 5G in 2020 as well as beyond 2020. As such, not all the requirements will need to be satisfied in 2020. Nevertheless, the 5G technology baseline should be designed so that it allows all these requirements to be satisfied at some point in or beyond 2020. The exact requirement set for the first release of 5G (addressing deployments around 2020) will be the subject of a further prioritization exercise for NGMN operators in close cooperation with industry.

4.1 User Experience

User experience requirements address the end user's experience when consuming one or more services. User experience will have to be managed in highly heterogeneous environments and under different user scenarios/contexts.

4.1.1 Consistent User Experience

The 5G system should be able to deliver a consistent user experience over time for a given service everywhere the service is offered. Consistent user experience is defined by service-dependent minimum KPIs (e.g. data rate, latency) being met over the service coverage area, with a level of variation configurable by the operator. These service-dependent KPIs are for further study and evaluation, within the NGMN program and elsewhere across the ecosystem, as appropriate. A consistent user experience across time and space depends obviously on the technology performance and capabilities, and on the operator deployment. The requirements address only the technology performance and capabilities.

4.1.2 User Experienced Data Rate

Data rate requirements are expressed in terms of user experienced data rate, measured in bit/s at the application layer. The required user experienced data rate should be available in at least 95% of the locations (including at the cell-edge) for at least 95% of the time within the considered environment. The user experienced data rate requirement depends on the targeted application/use case. It is set as the minimum user experienced data rate required for the user to get a quality experience of the targeted application/use case.
Use case specific user experienced data rates up to 1 Gb/s should be supported in some specific environments, like indoor offices, while at least 50 Mb/s shall be available everywhere cost-effectively. Use case specific user experienced data rate requirements are specified in Table 1.

4.1.3 Latency

When considering latency requirements, the following metrics are considered:

- E2E Latency: Measures the duration between the transmission of a small data packet from the application layer at the source node and the successful reception at the application layer at the destination node plus the equivalent time needed to carry the response back.
- User Plane Latency: Measures the time it takes to transfer a small data packet from user terminal to the Layer 2 / Layer 3 interface of the 5G system destination node, plus the equivalent time needed to carry the response back.

The E2E latency is the latency perceived by the end user. It accounts for the time needed for the data packet to cross all the nodes up to the application server and back, which includes nodes of the 5G system and nodes potentially outside the 5G system. In contrast, the user plane latency is limited to the 5G system only. Both latency metrics approximately coincide when the application server is located within the 5G system. In the latter case, the latency is minimised when the application server is co-located with a radio node, e.g., the radio base station or another user terminal (for the case of device-to-device, D2D, communication). As a result, the requirements on minimum latency are expressed in terms of E2E latency.

The 5G system should be able to provide 10 ms E2E latency in general and 1 ms E2E latency for the use cases which require extremely low latency. Note these latency targets assume the application layer processing time is negligible to the delay introduced by transport and switching. Use case specific E2E latency requirements are specified in Table 1.

The 5G system should also give the end user the perception of being always connected. The establishment of the initial access to the network (or status change from idle state to connected) should then be instantaneous from the end user perspective.

4.1.4 Mobility

Mobility refers to the system's ability to provide seamless service experience to users that are moving. In addition to mobile users, the identified 5G use cases show that 5G networks will have to support an increasingly large segment of static and nomadic users/devices. 5G solutions therefore should not assume mobility support for all devices and services but rather provide mobility on demand only to those devices and services that need it. In other words, mobility on-demand should be supported, ranging from very high mobility, such as high-speed trains/airplanes, to low mobility or stationary devices such as smart meters.

The mobility requirements are expressed in terms of the relative speed between the user and the network edge, at which consistent user experience should be ensured (see Consistent User Experience requirement). Use case specific mobility requirements are specified in Table 1.

4.1.5 User Experience KPI's

Table 1: User Experience Requirements

Use case category	User Experienced Data Rate	E2E Latency	Mobility
Broadband access in dense areas	DL: 300 Mbps UL: 50 Mbps	10 ms	On demand, 0-100 km/h
Indoor ultra-high broadband access	DL: 1 Gbps, UL: 500 Mbps	10 ms	Pedestrian
Broadband access in a crowd	DL: 25 Mbps UL: 50 Mbps	10 ms	Pedestrian
50+ Mbps everywhere	DL: 50 Mbps UL: 25 Mbps	10 ms	0-120 km/h
Ultra-low cost broadband access for low ARPU areas	DL: 10 Mbps UL: 10 Mbps	50 ms	on demand: 0-50 km/h
Mobile broadband in vehicles (cars, trains)	DL: 50 Mbps UL: 25 Mbps	10 ms	On demand, up to 500 km/h
Airplanes connectivity	DL: 15 Mbps per user UL: 7.5 Mbps per user	10 ms	Up to 1000 km/h
Massive low-cost/long-range/low-power MTC	Low (typically 1-100 kbps)	Seconds to hours	on demand: 0-500 km/h
Broadband MTC	See the requirements for the Broadband access in dense areas and 50+Mbps everywhere categories		
Ultra-low latency	DL: 50 Mbps UL: 25 Mbps	<1 ms	Pedestrian
Resilience and traffic surge	DL: 0.1-1 Mbps UL: 0.1-1 Mbps	Regular communication: not critical	0-120 km/h
Ultra-high reliability & Ultra-low latency	DL: From 50 kbps to 10 Mbps; UL: From a few bps to 10 Mbps	1 ms	on demand: 0-500 km/h
Ultra-high availability & reliability	DL: 10 Mbps UL: 10 Mbps	10 ms	On demand, 0-500 km/h
Broadcast like services	DL: Up to 200 Mbps UL: Modest (e.g. 500 kbps)	<100 ms	on demand: 0-500 km/h

4.2 System Performance

System performance requirements define the system capabilities needed to satisfy the variety and variability of users and use cases.

4.2.1 Connection Density

Up to several hundred thousand simultaneous active connections per square kilometre shall be supported for massive sensor deployments. Here, *active* means the devices are exchanging data with the network. Note this KPI assumes a single operator in the considered area.

Use case specific Connection Density requirements are specified in Table 2.

4.2.2 Traffic Density

The 5G network should be able to serve massive number of HTC and MTC devices. In the extreme cases:

- Data rates of several tens of Mb/s should be supported for tens of thousands of users in crowded areas, such as stadiums or open-air festivals.
- 1 Gb/s to be offered simultaneously to tens of workers in the same office floor.

Use case specific Traffic Density requirements are specified in Table 2.

Traffic Density measured in bit/s/m^2 is defined as the total amount of traffic exchanged by all devices over the considered area. The KPI requirement on the minimum Traffic Volume Density / Areal Capacity for a given use case is given by the product: [required user experienced data rate] x [required connection density]. For the sake of defining this KPI, a single operator is considered in the considered area.

4.2.3 Spectrum Efficiency

Spectrum efficiency should be significantly enhanced compared to 4G in order for the operators to sustain such huge traffic demands under spectrum constraints, while keeping the number of sites reasonable. Spectrum efficiency improvements should apply in both small and wide area cells, in both low and high frequency bands, in both high and low mobility scenarios.

In particular the average spectrum efficiency (measured in bit/s/Hz/cell) and the cell-edge spectrum efficiency (measured in bit/s/Hz/user) should be improved.

4.2.4 Coverage

The 5G technology should allow the data rates requirements to be achieved in rural areas with only the current grid of macro sites. The coverage requirement for other environments is for further study, within the NGMN program, and elsewhere, as appropriate.

4.2.5 Resource and Signalling Efficiency

Signalling efficiency should be enhanced, so that the related radio resource and energy consumption are minimised and justified by the application needs. More specifically, network function specific signalling (in particular related L1 control information) should only be transmitted when needed. In this context, UE capability handling should also be designed for network flexibility and scalability.

For certain IoT/MTC applications, additional measures should be considered to avoid a surge by volume in case a large number of devices attempt to access the network simultaneously.

4.2.6 System Performance KPI's

Table 2: System performance requirements

Use case category	Connection Density	Traffic Density
Broadband access in dense areas	200-2500 /km^2	DL: 750 Gbps / km2 UL: 125 Gbps / km2
Indoor ultra-high broadband access	75,000 / km^2 (75/1000 m^2 office)	DL: 15 Tbps/ km2 (15 Gbps / 1000 m2) UL: 2 Tbps / km2 (2 Gbps / 1000 m2)
Broadband access in a crowd	150,000 / km^2 (30.000 / stadium)	DL: 3.75 Tbps / km2 (DL: 0.75 Tbps / stadium) UL: 7.5 Tbps / km2 (1.5 Tbps / stadium)
50+ Mbps everywhere	400 / km^2 in suburban 100 / km^2 in rural	DL: 20 Gbps / km2 in suburban UL: 10 Gbps / km2 in suburban DL: 5 Gbps / km2 in rural UL: 2.5 Gbps / km2 in rural
Ultra-low cost broadband access for low ARPU areas	16 / km^2	16 Mbps / km^2
Mobile broadband in vehicles (cars, trains)	2000 / km^2 (500 active users per train x 4 trains, or 1 active user per car x 2000 cars)	DL: 100 Gbps / km^2 (25 Gbps per train, 50 Mbps per car) UL: 50 Gbps / km^2 (12.5 Gbps per train, 25 Mbps per car)
Airplanes connectivity	80 per plane 60 airplanes per 18,000 km^2	DL: 1.2 Gbps / plane UL: 600 Mbps / plane
Massive low-cost/long-range/low-power MTC	Up to 200,000 / km^2	Non critical
Broadband MTC	See the requirements for the Broadband access in dense areas and 50+Mbps everywhere categories	
Ultra-low latency	Not critical	Potentially high
Resilience and traffic surge	10,000 / km^2	Potentially high
Ultra-high reliability & Ultra-low latency* **(*) the reliability requirement for this category is described in Section 4.4.5**	Not critical	Potentially high
Ultra-high availability & reliability* **(*) the reliability requirement for this category is described in Section** 4.4.5	Not critical	Potentially high
Broadcast like services	Not relevant	Not relevant

4.3 Device Requirements

Smart devices in the 5G era will grow in capability and complexity as both the hardware and software, and particularly the operating system will continue to evolve. They may also in some cases become active relays to other devices, or support network controlled device-to-device communication.

4.3.1 Operator Control Capabilities on Devices

5G terminals should have a high degree of programmability and configurability by the network, for example in terms of terminal capabilities, access technology used, transport protocol used and certain lower layer functions (e.g. error control schemes). This will enable efficient logical division for different services (slicing) while removing dependency on terminal type.

In particular, flexible and dynamic UE capability handling should be assured. This would allow the network or the UE to choose one of the profiles depending on QoS needs, radio node capability and/or radio conditions.

The 5G devices should provide the capability to operators to check the hardware and software platform configuration over the air, the capability to update the smart device's operating system over the air, and the ability to diagnose the malfunction of devices or malware in smart device plus the ability to fix the problems or update device software that affect end user experience or overall network performance.

Operators shall be able to retrieve network as well as service-related performance data (e.g., voice call drops, handover failure, network registration failure, instantaneous throughput) from the UE in order to collect information on real-life operation and use them as an input for service experience optimization and customer care.

4.3.2 Multi-Band-Multi-Mode Support in Devices

To enable true global roaming capability, smart devices should be able to support multiple bands as well as multiple modes (TDD/FDD/mixed). Note that IoT/MTC devices which are stationary may not require multiple bands/modes.

Furthermore, to achieve the high data rates, devices should be able to use multiple bands simultaneously, without impacting the single band performance or network performance. 5G terminals shall support aggregation of data flows from different technologies and carriers.

4.3.3 Device Power Efficiency

Battery life shall be significantly increased: at least 3 days for a smartphone, and up to 15 years for a low-cost MTC device.

4.3.4 Resource and Signalling Efficiency

At the device side, the resource and signalling efficiency requirement is even more crucial as frequent signalling has a significant impact on the battery life. The same requirement as in Section 4.2.5 applies for the device.

4.4 Enhanced Services

In the 5G environment, services will be developed using a rich set of network and value enabling capabilities as outlined in Chapter 3. Enhanced services will be characterized by a high level of security, experience and features.

4.4.1 Connectivity Transparency

Connectivity transparency is a key requirement for delivering consistent experience in a highly heterogeneous environment.

5G may involve a combination of radio access technologies (RATs). In addition, given that 3GPP LTE / LTE-Advanced is likely to further evolve within the 5G era, both new RATs and the LTE RAT may be accessible to 5G user terminals.

It is expected that a terminal may be connected to several RATs (including both new RATs and LTE) at a given instant, potentially via carrier aggregation, or by layer 2 (or higher) bandwidth aggregation mechanisms. This combination of RATs may involve also non-3GPP RATs, e.g., IEEE 802.11ax (High Efficiency Wi-Fi).

Each RAT will naturally be deployed from several radio access points; potentially comprising of both high transmit-power (macro type) and low transmit-power (e.g., micro, pico, femto) access points.

The connectivity transparency refers to the following requirements:

- The user application should be always connected to the RAT or combination of RATs and/or access point (or other user equipment in case of D2D) or combination of access points providing the best user experience without any user intervention (context-awareness);
- Further, the requirement above should be achieved in a seamless way from a user perspective. By defining the service interruption time as the time during which the user is not able to receive any user plane data, including inter-system authentication time, this requires:
  - Inter-RAT mobility service interruption time, including between 3GPP and non-3GPP RATs, shall be possible to be unnoticeable by the user (possibly depending on the user subscription).
  - Intra-RAT mobility service interruption time shall be possible to be unnoticeable by the user (possibly depending on the user subscription).
  - Seamless inter-system authentication, including between 3GPP and non-3GPP RATs.
- From the network perspective, the network shall be able to control the access points (or other user equipment in case of D2D) and RATs the user device will connect to, based on operator preferences and user’s subscription;
- In addition, 5G should provide new and more efficient connection management functionalities without the need to use legacy connection functions (e.g., Access Network Discovery and Selection Function - ANDSF).

4.4.2 Location

Contextual information is important for delivering instant and personalised services. Location is one of the most important contextual attributes. In 5G, network based positioning in three-dimensional space should be supported, with accuracy from 10 m to <1 m at 80% of occasions, and better than 1 m for indoor deployments. Tracking of high speed devices will be required to provide this location accuracy in a real-time manner.

5G network based localization should be able to cooperate with other/external techniques (e.g. with capability to pull data from partner sources) to further improve accuracy. The overall cost of network-assisted localization should be comparable to or lower than the current external means (e.g. satellite systems) or 4G solutions to acquire the location information.

On top of the accuracy requirement, the 5G system should enable the exposure of location information by the definition of an API that can be used for the development of location based services.

4.4.3 Security

Security has been one of the fundamental capabilities operators provide to their customers. 5G will support a wide range of applications and environments, from human-based to machine-based communication, and thus it should be able to deal with a huge amount of sensitive data that need to be protected against unauthorized access, use, disruption, modification, inspection, attack, etc. Moreover, since 5G should be capable to offer services for critical sectors such as Public Safety, eHealth, and utilities, the importance of providing a comprehensive set of features guaranteeing a high level of security is a core requirement for 5G systems. Therefore, 5G should be designed to provide more options beyond node-to-node and end-to-end security available in today's mobile systems, in order to protect users' data, create new business opportunities and prevent or mitigate any possible cyber security attack.

Subscriber Authentication

Similarly to 4G, a strong 5G authentication will represent a robust platform upon which operators can develop single-sign-on services: the 5G network operator, acting as an Identity Provider, could thus be responsible for users' identity authenticity towards external partners, providing transparent identification and seamless authentication to Application Services on behalf of the user. The subscriber's identity together with secret data allowing the access to a given network shall be stored in a secured physical entity (e.g. similar to the current UICC). The data necessary to access an operator network remain the sole ownership of the operator running this network.

The system shall offer the capability to protect 5G customers from common security threats (e.g., impersonation, traffic eavesdropping, etc.,) thus increasing the level of trust that is associated to their network subscribers' identity. Also, design of security solutions (e.g. key exchange/derivation protocols upon handover or when interworking with other RATs) should provide better secrecy than 4G without sacrificing efficiency.

User Privacy

The 5G system must provide security mechanism for protection of a variety of trusted information regarding human as well as machine-users (e.g., identity, subscribed services, location/presence information, mobility patterns, network usage behaviour, commonly invoked applications, etc.).

Beyond Hop-by-Hop Security

While radio bearer security may still be worthwhile, it is also useful to consider whether the 5G architecture can create additional business value by facilitating bearer-independent (e.g., higher layer) security, and extending to servers on the internet, or extending to device-to-device communications. Any mechanism conceived to realize such bearer-independent security should also be compliant to lawful interception obligations when these are required. Similarly, security mechanisms are needed to fight growing inter-operator fraud and misuse of international signalling networks. 5G roaming signalling protocols must enable the home network to verify that a user is really attached to a serving network that claims it is.

Network Security

With the massive penetration of IP protocols for control and user plane in all network functions, with the diffusion of low-cost MTC device or open OS smartphones where mobile malware could be easily propagated, the operator's 5G core and radio networks could become more vulnerable.

The following requirements highlight areas for improvements, with respect to LTE / LTE-Advanced (4G) security:

- Improve resilience and availability of the network against signalling based threats, including overload caused maliciously or unexpectedly
- Specific security design for use cases which require extremely low latency (including the latency of initiating communications)
- Comply with security requirements that are defined in 4G 3GPP standards. This will apply especially to a virtualized implementation of the network (virtual appliance, hypervisor)
- In the context of Public Safety and Mission Critical Communications, it is expected that 5G technology will allow reduction of cost and improvement of functionality of these networks. Besides supporting emergency communications, the 5G commercial system should be able to provide basic security functions in emergency situations, when part of the network infrastructure, including the security infrastructure, may be destroyed or inaccessible. The security services provided should be able to provide protection against malicious attacks that may intend to disrupt the network operation and allow the secure implementation and deployment of essential infrastructure.

From a purely radio access perspective, the following requirements emerge:

- Improve system robustness against smart jamming attacks of the radio signals and channels
- Improve security of 5G small cell nodes, taking into consideration their geographical distribution and their easy accessibility.

4.4.4 Resilience and High Availability

Resilience and high availability will be essential to ensure minimal service is available to critical infrastructures or service providers in case of disaster. Also, 5G networks will increasingly be used as the primary means for emergency communication and Public Safety for day to day operations.

The network availability is characterized by its availability rate X, defined as follows: the network is available for the targeted communication in X% of the locations where the network is deployed and X% of the time.
5G should enable 99.999% network availability, including robustness against climatic events and guaranteed services at low energy consumption for critical infrastructures (e.g., hospitals, network management). The level of network availability to be effectively provided is up to the operator.

Resilience, i.e. the capability of the network to recover from failures, will be an important feature to maintain high availability rates. In particular, remote (self-)healing of equipment should be possible.

4.4.5 Reliability

The reliability of a communication is characterized by its reliability rate, defined as follows: the amount of sent packets successfully delivered to the destination within the time constraint required by the targeted service, divided by the total number of sent packets. Note that the reliability rate is evaluated only when the network is available.

The reliability rate depends on the service and use case. The 5G technology should allow high reliability rates of 99.999%, or higher for the use cases that demand it, in particular those under the ultra-high reliability and ultra-low latency use cases category. For use cases for which reliability may be less an issue, e.g. some non-delay critical MTC use cases, the reliability rate may be 99% or even lower depending on the associated trade-off needs.

4.5 New Business Models

It is an essential requirement that 5G provides a future-proof technology platform allowing the evolution of existing business models in both retail and wholesale offerings. Furthermore, it should open up opportunities to create completely new business models without having an impact on network architecture. For network operators, the capability to evolve and enable new business models should be supported in a cost efficient manner, without having architectural impact. Using 5G networks, third party service providers should be able to offer their services in a very short time-to-market manner and based on mutual service level agreements, where the network will be delivering data using agreed network functions, capabilities and attributes.

5G should be designed from the beginning such that the network operator is able to create a large variety of relationships between its network infrastructure and the customer/service provider. With respect to the business model examples outlined earlier in the 5G vision, the following requirements are formulated.

4.5.1 Connectivity Providers

The connectivity provider business model applies for both retail and wholesale commercial relations and offerings. It assumes that the customer and service provider are decoupled from the physical infrastructure and they are offered no configurability, and a very low level of configurability, respectively. The 5G system should enhance the efficiency of this model by enabling the operator to configure the data flow to use only necessary functions in the network, on demand and in a programmable manner, in order to optimize operational and management costs. This requires modular network architecture, having the capability to be exposed to the 5G provisioning/configuration system.

4.5.2 Partner Service Provider and XaaS Asset Provider

Partner service provider and XaaS asset provider models refer to the 5G high-level requirement to allow creation of different levels of relationship between operators and application/service providers. Exploiting flexibility, 5G should be able to support different levels of abstraction and business models as known today (e.g., Infrastructure as a Service, Platform as a Service, Network as a Service) as well as allowing creation of completely new business models not foreseen at the time of writing this document. The key requirement is that Service providers should be able to configure and manage the service, while operators will have freedom to manage and evolve the network.

In this context, 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. Application/Service provider will then be able to use sub-set of the network capabilities in a flexible, configurable and programmable manner, and to use network resources depending on their service preference.

These exposed in-networking capabilities or information may include charging capabilities, authentication, mobility, reliability functions, mobile user's footprints, etc. For example, a 3rd party application can be charged for its user's traffic instead of charging to its users; operators can provide (regularly or on-demand) to an intelligent traffic management application the number of UEs presence in a certain area without violating user's privacy. Some radio information can also be exposed, such as real-time loading, QoS, UE measurement report, mobility, signal strength, etc. Real-time QoS may for example allow video applications to adjust UL/DL video bit rates to improve streaming experience.

4.5.3 Network Sharing Model

Network sharing business models involves a relationship between the service provider and the operator, and between operators, in which their respectively owned physical network infrastructures are tightly coupled. The 5G system should provide methods and instruments for various network sharing schemes developed to maximise the overall synergies of network sharing agreements and to enable flexible business models and commercial relationships that potentially may change rapidly. This should also apply to the current network sharing models. With 5G, it should be possible to provide sufficient flexibility to accommodate the capacity needs of dynamically hosted operators, on a real-time basis (e.g., for capacity brokering architecture, where network resources are provided dynamically depending upon bids offered).

Technical capabilities shall include spectrum sharing or reuse, enhanced mobility techniques and enhanced controls for access network, access point, access node, and spectrum selection at an operator policy level.

4.6 Network Deployment, Operation and Management

Network deployment, operation and management requirements are important to ensure sustainability and performance of the network.

4.6.1 Cost Efficiency

It is fundamentally essential to ensure that operators will be able to provide 5G services in an economically sustainable way. 5G should be designed with the objective to minimise the Total Cost of Ownership (TCO) of the network infrastructure as well as operation and management, for any given service offering. In addition, the cost of the devices should be minimized to facilitate the service access for the customers.

4.6.2 Energy Efficiency

Energy efficiency of the networks is a key factor to minimize the TCO, along with the environmental footprint of networks. As such, it is a central design principle of 5G.

Energy efficiency is defined as the number of bits that can be transmitted per Joule of energy, where the energy is computed over the whole network, including potentially legacy cellular technologies, Radio access and Core networks, and data centres.

5G should support a 1,000 times traffic increase in the next 10 years timeframe, with an energy consumption by the whole network of only half that typically consumed by today's networks. This leads to the requirement of an energy efficiency increase of x2000 in the next 10 years timeframe.

Every effort should be made to obtain the energy gain without degrading the performance, but the technology should allow native flexibility for the operator to configure trade-off between energy efficiency versus performance where justified.

4.6.3 Ease of Innovation and Upgrade

5G should provide efficient, flexible and fast ability for introduction of new services and future technical evolutions. In particular, building a new service should be much faster in 5G than in previous systems. The introduction of future technical evolutions includes introducing new features, as well as new access technologies, if needed to accommodate the future market needs. Such a flexibility requires an overall system architecture that allows any access technology (fixed and/or radio) to be connected to the same core network, including technologies unknown at the time of initial 5G design. Similarly, it should be possible to have innovation in the core network and the management system with minimal impact on the UE and access network.

4.6.4 Ease of Deployment

The 5G system should allow reusing or upgrading existing network infrastructures.
The 5G system should reduce the complexity of the tasks of planning, configuration and optimization of the whole system. In particular, the 5G system should allow for easy deployment and management of massive small cells with features, like plug and play, self-configuration, optimization and healing.

Flexible network deployment/topology should be possible to address highly diversified service requirements, such as high speed train, and super high local traffic sharing, etc. Network deployments should also be able to adapt to diversified network configurations, such as ideal/non-ideal, fixed/wireless backhaul/fronthaul, etc. The 5G system should fully enable low-cost and easy deployment by means of commodity hardware/software platform.

4.6.5 Flexibility and Scalability

One of the key characteristics of 5G networks will be the support of an extremely high variety of requirements in connection properties and attributes, driven by coexistence of very different use cases.

In the sense of connection attributes provided to the end user, 5G should enable openness and multi-vendor capability at all levels and introduce modular provisioning concept. It means that the key connection attributes (e.g., mobility, security & privacy, reliability, bandwidth, latency, etc.,) should be enabled/disabled/modified and controlled in a programmable and switchable manner depending on particular use-case and associated policy defined by Operator.

Strong spectrum agility should be possible, for a dynamic use of spectrum by different RAN technologies, depending on regulatory requirements.

In order to assure maximum flexibility and scalability during the technology lifecycle, 5G system design should adopt functional split of network domains as well as network elements:

- Core and RAN network domains should be functionally decoupled to create a radio technology agnostic architecture, where introduction and connection of new radio technology will be possible in a plug & play manner (see also Section 4.6.3).
- HW and SW functions of network elements shall be decoupled in all network domains.
- Changes and enhancements to one equipment/network domain should not mandate changes/enhancements to the other. This should allow cost- and effort-efficient upgrade path providing operators to leverage significant investments into existing operational infrastructure and maximize its utilization.
- Real-time and on demand network configuration and automated optimization should provide flexible and cost efficient network operation. In order to maximize utilization efficiency of available network resources, it should be possible to dynamically and freely relocate network resources depending on current and local needs, under full control of the operator.

4.6.6 Fixed-Mobile Convergence

The 5G system should support fixed and mobile convergence, in order to:

- Ensure a seamless customer experience within the fixed and mobile domains (e.g., a unified user authentication).
- Allow the operator to process a customer independently of his access type for authentication and billing, via a unified customer data base and information system across the fixed and mobile domains.

4.6.7 Operations Awareness

The 5G system should be able to access, monitor and process various pieces of information in order to optimise its operation, in particular:

- Instantaneous network conditions to optimally connect and route user traffic in a dynamic manner;

- Service / application traffic characteristics, so that operator policy/control can be enforced in a timely manner to optimize traffic flows.

In addition, the operator of the 5G system should be able to securely collect information that can enhance user experience and service experience (e.g. speed, location) via data analytics. Security for all the subscriber information collected by the 5G system should be provided as described in section 4.4.3 Security.

4.6.8 Operation Efficiency

Operation efficiency is a key component to reduce costs and energy consumption in the network. The 5G system is expected to introduce new challenges in the operational processes. The RAN deployment, characterized by e.g., increased base station density, higher frequency spectrum and heterogeneous environment, as well as the coexistence with the legacy deployment, will increase the complexity of network management process. In 5G, there is no more real border between backhaul and mobile core network. This network management complexity may impact all network domains and the way to manage the supported services.

This motivates the following requirements:

- The complexity of Management and Operations and associated OPEX and CAPEX shall be significantly reduced in comparison with today's environment.
- Flexible, programmable and real time network and service management processes, relying on autonomic/self-management functions (self-configuring, self-diagnosing, self-healing and self-optimising network) shall be supported in a harmonized way because they share common objectives.
- Autonomic/self-management functions shall guarantee "trust & confidence", stable and self-coordinated behaviour within a single network domain (RAN, backhaul or core network) and across different network domains (RAN/CN, multi-vendor, multi-RAT).
- Autonomic/self-management functions shall be provided both at management plane and at control plane level and shall support a flexible architecture (e.g., centralized and/or distributed) to cope with different use cases.
- It shall be possible to trigger measurement campaigns at network level (e.g., for a given cell or group of cells) with support of targeted UEs on user consent-basis. In addition, user/application level QoS/ QoE monitoring capability shall be supported by UE and network and controlled at network / service management level i.e., to extend the monitoring to the application level in order to be able to introduce metrics that can provide a characterization of the user experience (e.g., for a video streaming service).
- All needed management interfaces in 5G network shall be deployed using a new paradigm of standard interfaces (APIs) at various layers (e.g., Southbound and Northbound) with support of open source community. Open-source based solutions have to be considered.
- 5G operations / management framework must be able to expose open management APIs to allow partners to exchange management information. Those capabilities must cover fulfilment (ordering, provisioning, service and resource activation, service and resource inventory) and assurance (performance management, alarm management, threshold management).
- The above requirements shall be met at the border between shared RAN and unshared RANs (of sharing operators) in a 5G "RAN Sharing" management model.
- The above requirements shall be met in the following business scenarios: "Unshared" 5G Network (B2C) and "Sharing Network" (B2B2C) - at mobile backhaul, at RAN level, at mobile transport network level and at mobile core network level.

- Within B2B2C business scenarios, a governance / policy shall be clearly defined from a Management & Operations perspective. This means actors & roles and related duties & rights as well as responsibilities attached to participating Operations Support Systems (OSSs) shall be well defined. Policy, SLAs and associated APIs and the format of data exchanged between those actors shall be defined in a common (and possibly standardized) way.

4.6.9 Ultra Low-cost Networks for Very Low-ARPU Areas

Although central for any market, cost efficiency is even more important for very low Average Revenue Per User (ARPU) areas, e.g., rural areas with very low densities of populations, and rural/suburban areas not yet connected to the Internet because of economic constraints.

A flavour of 5G is expected to be flexible enough to be deployed under ultra-low cost requirements to offer decent Internet access to the remaining inhabitants on Earth who do not have access to the internet. Bringing connectivity to such areas in an economically sustainable way requires ultra-low cost network infrastructures, ultra-low cost devices, and ultra-low cost operation and maintenance.

In addition to minimizing the costs of the full-fledged technology, 5G therefore needs to offer options and possibilities for ultra-low cost deployments tailored for very low ARPU areas.

Features of lower importance for low cost deployments include:

- Lower availability: a high availability typically requires some redundancy of equipment. For ultra-low cost networks the availability rate requirement can be lowered.
- Lower peak rates: Lowering peak rates can be enabled by removing features like higher order modulation, MIMO configurations, carrier aggregation support, etc.
- Mobility limitations: Inter RAT mobility functions can be removed. Intra-RAT mobility functions can be simplified if it helps decreasing the cost of infrastructure and devices.
- Restricted periods of service: at peak hours, when large numbers of users attempt to connect to the network but serving all users simultaneously is not economically viable, connectivity can be shared in time between the users. This leads to periods of service unavailability from the user perspective, but allows reasonably costly network dimensioning.
- Energy-level dependent base stations activity: When the energy level of a base station operating off the grid (e.g., on battery) reaches a certain threshold, the base station may enter an energy saving mode where the service to regular users may be degraded (e.g., the transmit power may be reduced) or even shut down in order to save energy for public safety services.
- Restricted areas of service: Consistency of user experience across a wide territory is not mandatory. Only minimum communication services may be available everywhere, with higher bandwidth being available only in some areas (e.g., where population is present).

4.6.10 Ultra Low-Cost Networks for Very Low-ARPU MTC Services

In addition, some large-scale MTC services will generate only a very low ARPU, e.g., 100 times lower than current human users. The 5G system should enable economically viable deployments to address this market, by providing a sufficiently low associated TCO.

Features of lower importance for low cost deployments in this case include the following ones from the list in the previous section:

- Lower peak rates;
- Mobility limitations;
- Restricted periods of service.

5. TECHNOLOGY AND ARCHITECTURE

5.1 Analysis

When the baseline 4G system (which in here is considered to be 3GPP Release-12) is compared against the 5G requirements, improvements are needed in three dimensions, namely, network capabilities, enablers for operational sustainability, and enablers for business agility.

The ***network capabilities*** of 3GPP Release-12 fall short of the NGMN requirements in a number of areas, as illustrated in Table 3.

Table 3: Overview of selected 3GPP Release-12 network capabilities/ typical implementation and foreseen improvements necessary to meet the NGMN requirements

Attribute	3GPP Release-12 capability	Improvement needed to meet NGMN requirements	Remarks
Data rate (per user)	Up to 100 Mb/s on average Peaks of 600 Mb/s (Cat 11/12)	> 10X expected on average and peak rates > 100X expected on cell edge	
End-to-end latency	10 ms for two-way RAN (pre-scheduled) Typically, up to 50 ms end-to-end if other factors are considered (e.g., transmission, CN, internet, proxy servers)	> 10X (smaller)	Technology should allow operators to optimize topology to achieve 1 ms end-to-end.
Mobility	Functional up to 350 km/h (for certain bands up to 500 km/h) No support for civil aviation	> 1.5X	Functional in 5G means sustained service quality for the considered use case. 5G in addition should support civil aviation use case.
Spectral efficiency	DL: 0.074 – 6.1 b/s/Hz UL: 0.07 – 4.3 b/s/Hz depending on cell edge or average, deployment scenario, and FDD or TDD	Pushing the envelope for substantial increase	Requirements should be specified by NGMN operators jointly with the industry in due course.
Connection density	Typically ~2,000 active users/km^2	> 100X	

Current networks also lack sufficient enablers for ***operational sustainability***. Despite the development of self-organizing network (SON) functions for LTE, base station configuration, fault identification, troubleshooting and ongoing optimizations all require human resources and site visits in some cases, which is costly in both time and monetary resources. In addition, the core network architecture specifies many entities, such as P-GW, S-GW, MME, PCRF, OCS, OFCS, ANDSF, etc., which makes it challenging to deploy and manage. With such complexity, the network becomes expensive and costly to scale with growing number of devices and growing traffic volume. The current network hardware is also mostly specialized vendor-specific hardware which limits the flexibility in functional use. Also, the lack of inbuilt monitoring tools means that external probes are required, which adds to the cost of running the network. In addition, the energy efficiency of the networks today is not scalable to support the data rates and capacity envisaged for 5G.

Enablers for improved ***business agility*** are also required. The 4G system is designed from the ground up to primarily support mobile broadband services. As a result, the signalling, management and accounting procedures (e.g., identity, initial access, AAA, etc.) are not well suited for the provisioning of services that do not cleanly fit in the mobile broadband category. Moreover, the interfaces to other data networks (other than 3GPP networks) and to the services layer are not sufficiently specified. As a result, adding new capabilities to launch new services usually require expensive proprietary solutions. A further consequence of this is that the types of business models (including potential partners) that are efficiently supported are limited.

From the above, it is clear that substantial improvements are needed in all three dimensions to bring the current state of the art system on par with the requirements. Existing systems (e.g., 3GPP Release-12, IEEE 802.11) are continuously evolving in terms of standardization, implementation and deployment. Several ongoing trends are identified in Figure 6. These are expected to help the existing systems to improve in the three dimensions. Nevertheless, incremental evolution of 4G systems alone is not expected to be sufficient to address all the shortfalls. Thus, a 5G system is required that introduces some fundamentally new technologies and paradigms to complement ongoing evolutionary trends.

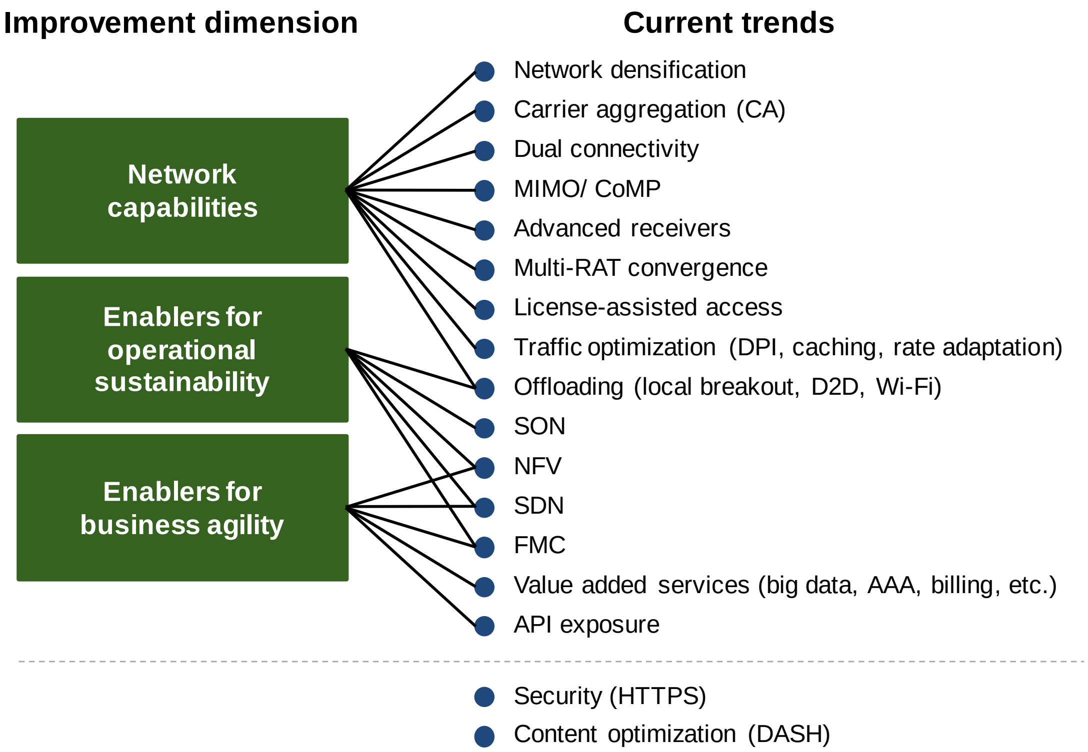


Figure 6: Ongoing technology trends

5.2 Technology Candidates

Some of the technologies depicted in Figure 6 are expected to be at an advanced stage of maturity before 5G deployment. Thus, 5G networks should be designed to leverage them. Besides these, new technologies currently in their early stages of development, such as (but not limited to) massive MIMO and full duplex, should be considered. A detailed initial list of technologies that are visible to NGMN are provided in the Annex, together with their benefits, areas of applicability, maturity level, requirements addressed, and foreseen issues to be investigated. Note that the list may not be exhaustive and any emerging technologies available in due time for 5G should be considered.[1] These technologies together could potentially address the majority of the requirements (see Section 14 in the Annex). Nevertheless, the extent and the degree to which they will address the requirements need to be carefully studied by NGMN operators jointly with the industry.

To illustrate, the network capacity of a radio network depends on the spectral efficiency, spectrum bandwidth and cell density. Shannon's capacity equation is a fundamental constraint on the spectral efficiency performance and implies logarithmic dependence of channel capacity on the signal-to-interference and noise ratio (SINR). Interference mitigation and coordination techniques could improve the effective SINR, especially at low SINR regions, thereby improving the system spectral efficiency. Massive MIMO could also improve the SINR through narrow beamforming, pushing the system closer to a noise limited environment. If the system could realize high SINR regions in a wider area through use of such technologies, advanced coding and modulation schemes may provide more effective gains. Furthermore, flexible and full duplex, as well as schemes which reduce the amount of guard bands and overhead, may improve the overall spectral efficiency. However, given the theoretical limit and technologies, it seems unrealistic to assume order of magnitude improvements.

The use of new spectrum in licensed and license-exempt bands should enable increases in data rate and capacity. However, substantial amounts of new spectrum are only likely to be found in higher frequency bands (e.g., millimetre waves (mmW)) with expected propagation restrictions, particularly when needing to penetrate into buildings from the outside. Massive MIMO is well suited for higher frequency bands, but the cost-effectiveness may constrain the technology's application. New designs will be needed to make practical antenna form factors for massive MIMO at lower carrier frequencies. At higher frequencies where a large number of antenna elements can be fitted within practical form factors, the optimal transceiver implementation that balances beamforming performance (e.g., phase calibration) with cost is still a subject of research.

Given the constraints on spectrum efficiency and higher frequency band deployments discussed above, network site densification will be an important approach to deliver substantial data rate and capacity gains, particularly as it also supports the use of higher frequency spectrum. Nevertheless, both wider spectrum bandwidth and massive MIMO impose higher fronthaul capacity requirements for centralized approaches to densification (i.e., use of RRUs or DAS) which will affect the technical and economic viability.

As a further example, virtualization together with flexible network function definition and allocation is an important enabler for network adaptability. However, the best approach to redesigning network functionalities to leverage the benefits of NFV is not immediately obvious. Simply virtualizing existing network nodes (e.g., gateways, MMEs) may make them cheaper to implement but will not reduce the network complexity or provide the needed adaptability to specific use cases. Opting for much finer granularity of functions (e.g., mobility management, access authorization, encryption) will require interfaces between all possible function combinations to be defined to foster interoperability among functions from different vendors – not necessarily as protocols, but potentially as software interfaces (e.g., service-oriented architecture (SOA) - based). This could result in management complexity and interoperability testing efforts, as well as market fragmentation. On the other hand, defining specialized blocks comprising several individual functions (e.g., to support machine communications) could increase business agility, but will lead to duplication of functionalities across such specialized blocks and potentially increased costs.

Technologies may have adverse or complementary effects on each other. Hence, finding the right combination of technologies to address all the NGMN requirements becomes very challenging. Fortunately, not all requirements need to be met at the same time for different use cases. This makes it

[1] For example, satellite communications may have relevance as a component in 5G to address some requirements.

possible to apply different combinations of technologies to different use cases. Nevertheless, this flexibility should not come at the expense of increased costs and complexity.

5.3 5G Design Principles

Given the requirements stipulated earlier, and considering emerging technology trends, NGMN believes that the 5G system should be designed based on the design principles illustrated in Figure 7. These principles are further elaborated below.

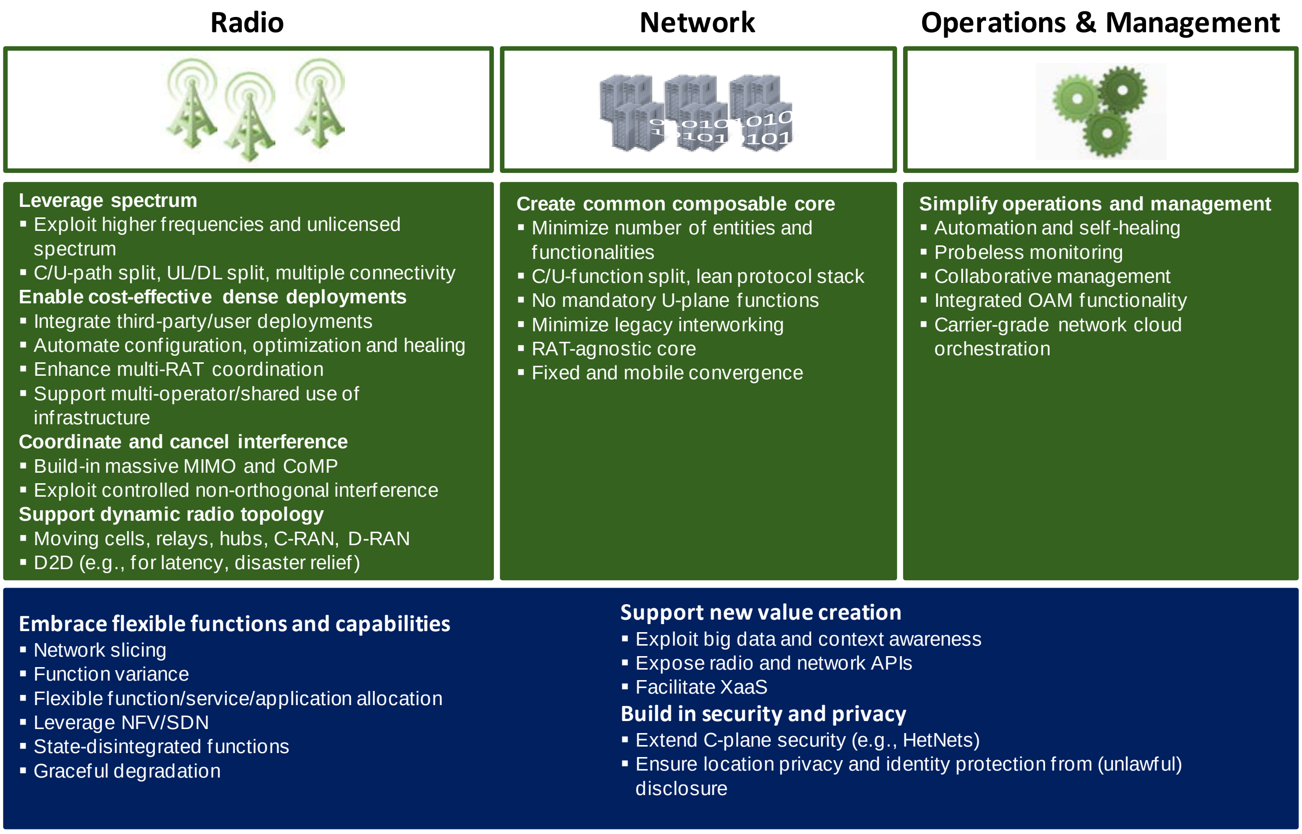


Figure 7: 5G design principles

5.3.1 Radio

Leverage spectrum – Higher frequencies (e.g., centimetre and millimetre waves) and licence-exempt spectrum should be exploited to complement endeavours to use any spare bandwidth at lower frequencies and as a complement to the available exclusively licensed mobile spectrum resource. Due to different properties of different spectrum, concepts such as C/U-plane path split and UL/DL split should be employed to optimize the use of various spectra. This implies that simultaneous connections to multiple access points need to be supported.

To optimize the spectrum use depending on the traffic demand, flexible duplex should be facilitated by design, e.g., via a unified frame structure. In addition, full duplex should be applied where feasible, to resolve issues around FDD (e.g., guard bands) and TDD (e.g., guard time, synchronization). Even if implementation technologies limit the achievable performance by 2020, protocols should be designed to support flexible and full duplex from the beginning, if advances in implementation technologies are foreseen.

In addition, the RF capabilities of devices must be improved to take full advantage of different spectrum opportunities whilst maintaining power-efficient large bandwidth operation without desensitization.

Enable cost-effective dense deployments – With extreme densification, cell planning and coordinated deployment will become increasingly difficult, since deployment becomes 3D and site negotiations will become more difficult, resulting in sub-optimal sites. To make densification economically viable, new deployment models, such as integration of third-party/ user deployments as well as multi-operator/ shared deployments are necessary. The system should be able to cope with unplanned, chaotic deployments and unexpected interference, drawing out maximum performance even given such deployments. The network should hence be designed to adapt to the availability of different types of backhaul and fronthaul and to have automated configuration, optimization and healing capabilities. This includes the ability to self-manage interference and load balancing.

Enhanced multi-layer and multi-RAT coordination, as well as dynamic/ fast switching between frequencies, cells, beams and RATs are necessary to ensure seamless user experience while mobile in such dense deployments. To support such control, effective mechanisms to detect UE speed as well as direction of movement are needed.

In addition, to support these features under multi-vendor deployments, open interfaces between C- and U-plane functions will be required to allow the U-plane functions on different platforms to be consistently controlled by common C-plane functionality.

Coordinate and cancel interference – Massive MIMO and CoMP will be essential to improve the achievable SINR in the system, thereby improving QoS consistency and overall spectrum efficiency. Both massive MIMO and CoMP transmission rely on the availability of channel state information to realize their full potential. Thus, efficient mechanisms to obtain the necessary information must be considered in the design from the start. Considering the variety of different CoMP methods already proposed for LTE, ranging from coordinated scheduling to joint transmission, 5G should be designed to natively support the most effective techniques. The 5G network architecture should therefore support the flexible location of coordination functions depending upon the transport network capabilities, supporting a trade-off between the benefits of wide-area optimization in more central locations with the potentially detrimental impact of backhaul delay on resource allocation.

The 5G network must also be designed to exploit any feasible interference cancellation methods, such as non-orthogonal multiple access (NOMA) with advanced receivers, where they offer useful performance benefits.

Support dynamic radio topology – Devices should be connected through topologies that minimize battery consumption and signalling, without limiting their visibility and reachability by the network, when desired. Wearable devices could connect through a smartphone as well as directly to the network if the smartphone battery runs out. Extended use of tinted glass on vehicles as well as large scale sensor deployment makes hub devices highly relevant. In some cases D2D communications could be exploited to offload traffic from the network. Thus, the radio topology should be able to change dynamically based on the context. A unified frame design, together with radio topology-agnostic design of identities, authentication and mobility procedures, is essential to support this.

5.3.2 Core Network

Create common composable core – To support the diversity of use cases and requirements in a cost-effective manner, the system design should move away from the 4G monolithic design optimized for mobile broadband. In this regard, a rethink of models such as bearers, APNs, extensive tunnel aggregation and gateways is needed. In addition, the UE state machine and entities which store UE context should be revisited and redesigned. Mandatory functions should be stripped down to an absolute minimum, and C/U-plane functions should be clearly separated with open interfaces defined between them, so that they can be employed on demand.

To provide further simplification, legacy interworking must also be minimized, for example towards circuit switched domain in the 2G and 3G networks. A converged access-agnostic core (i.e., where identity, mobility, security, etc. are decoupled from the access technology), which integrates fixed and mobile core on an IP basis, should be the design goal.

5.3.3 End-to-End

Embrace flexible functions and capabilities – Network/ device functions and RAT configuration should be tailored for each use case, leveraging the NFV and SDN concepts. Thus, the network should support flexible composition of network functions, as well as, their flexible allocation and location. The network functions should be scalable such that capacity is provided when and where needed. Even when particular functions or nodes become unavailable, e.g., due to disaster events, the system should support graceful degradation instead of complete service interruption. To improve such robustness, state information should be split from functions and nodes, so that contexts could be easily relocated and restored even in failure events.

5G should aim to virtualize as many functions as possible, including the radio baseband processing. Although some functions may still run on non-virtualized platforms, e.g., to meet state-of-the-art performance targets, they should be programmable and configurable using C-plane functions according to SDN principles.

Support new value creation – 5G should make it possible to exploit the network to quickly and efficiently create new value added services and explore different business models and opportunities. For instance, big data and context awareness can be used to create new values for third-party and social use, e.g., for marketing, optimizing public transport, and city planning. Thus, the network design must make the collection, storage and processing of the necessary data simple and efficient.

To further benefit from a programmable network platform, appropriate APIs to various parts of the network should be exposed and standardized. This enables access by third-parties and fosters the realization of different XaaS business models. For example, the APIs could allow third-party access to agile service creation, network measurements, network traces and full configuration control of network functions to enable seamless configuration changes in real-time.

Build in security and privacy – Security is an essential value proposition of the 5G system and must be a fundamental part of the system design despite paradigm shifts like extreme densification, dynamic radio topology, and flexible function allocation. In particular, user location and identity must be protected from unlawful disclosure. Some 5G use cases require extremely low latency – including the latency of initiating communications. For these use cases, multiple-hop security, where intermediate nodes need to decrypt and re-encrypt data, should be avoided.

It should be noted that end-to-end security methods (e.g., SSL, VPN and HTTP 2.0) are increasingly prevalent and these provide the added benefit of protection outside of the 5G operator domain. This could give rise to unfortunate duplication of security functions both in the network and at the communication endpoints. Nonetheless, not all communications in future may receive sufficient end-to-end protection. Hence a flexible architecture could help to tailor the network security functions to suit the application.

5.3.4 Operations & Management

Simplify operations and management – Expanded network capabilities and flexible function allocation should not imply increased complexity on operations and management. Procedures should be automated as far as possible, with well-defined open interfaces to mitigate multi-vendor interworking problems as well as interoperability (roaming) issues. Use of dedicated monitoring tools should be avoided and network functions (software) should be embedded with monitoring capabilities. Big data analysis should drive network management from reactive to a predictive and proactive mode of operation. Carrier-grade network cloud orchestration is needed to ensure network availability and reliability.

5.4 5G Architecture

Based on the design principles, NGMN envisions an architecture that leverages the structural separation of hardware and software, as well as the programmability offered by SDN and NFV. As such, the 5G architecture is a native SDN/ NFV architecture covering aspects ranging from devices, (mobile/ fixed) infrastructure, network functions, value enabling capabilities and all the management functions to orchestrate the 5G system. APIs are provided on the relevant reference points to support multiple use cases, value creation and business models. This architecture is illustrated in Figure 8.

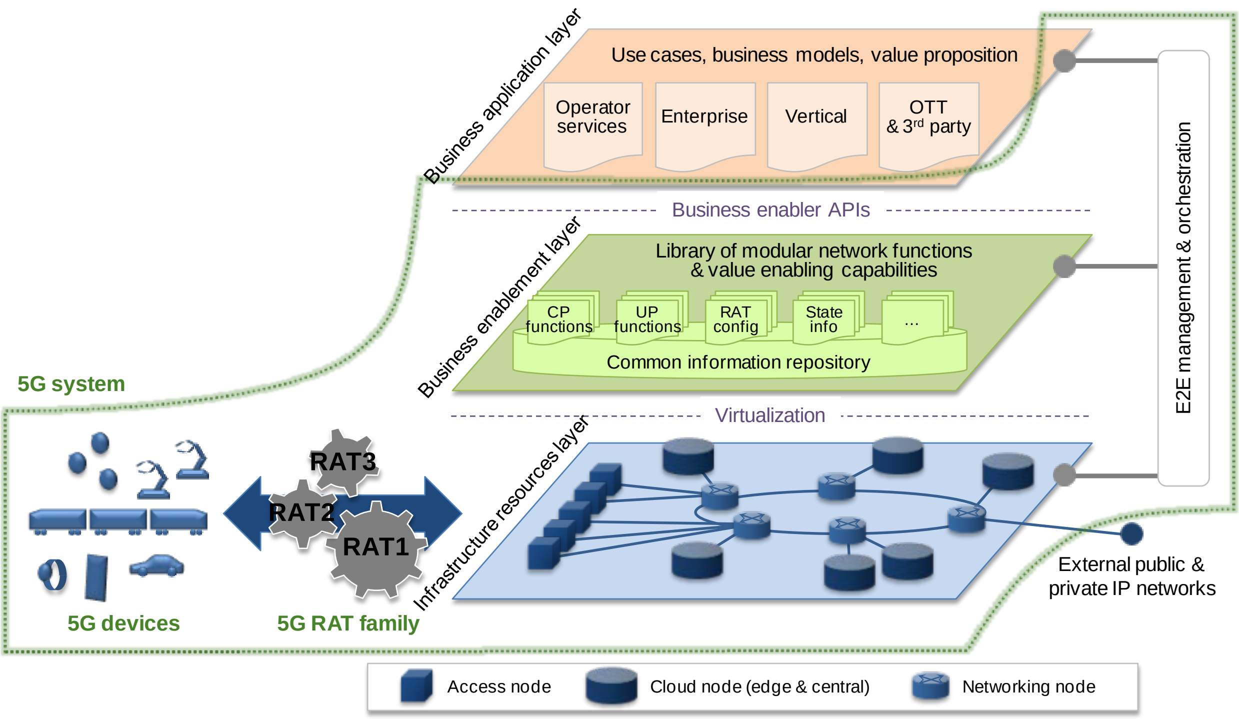


Figure 8: 5G Architecture

The architecture comprises three layers and an E2E management and orchestration entity.

The ***infrastructure resource layer*** consists of the physical resources of a fixed-mobile converged network, comprising access nodes, cloud nodes (which can be processing or storage resources), 5G devices (in the form of (smart) phones, wearables, CPEs, machine type modules and others), networking nodes and associated links. 5G devices may have multiple configurable capabilities and may act as a relay/ hub or a computing/ storage resource, depending on the context. Hence, 5G devices are also considered as part of the configurable infrastructure resource. The resources are exposed to higher layers and to the end-to-end management and orchestration entity through relevant APIs. Performance and status monitoring as well as configurations are intrinsic part of such an API.

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 realized by software modules that can be retrieved from the repository to the desired location, and a set of configuration parameters for certain parts of the network, e.g., radio access. 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. The different levels of performance and capabilities offered could be utilized to differentiate the network functionality much more than in today's networks (e.g., to offer as mobility function nomadic mobility, vehicular mobility, or aviation mobility, depending on specific needs).

The ***business application layer*** contains specific applications and services of the operator, enterprise, verticals or third parties that utilize the 5G network. The interface to the end-to-end management and orchestration entity allows, for example, to build dedicated network slices for an application, or to map an application to existing network slices.

The ***E2E management and orchestration entity*** is the contact point to translate the use cases and business models into actual network functions and slices. It defines the network slices for a given application scenario, chains the relevant modular network functions, assigns the relevant performance configurations, and finally maps all of this onto the infrastructure resources. It also manages scaling of the capacity of those functions as well as their geographic distribution. In certain business models, it could also possess capabilities to allow for third parties (e.g., MVNOs and verticals) to create and manage their own network slices, through APIs and XaaS principles. Due to the various tasks of the management and orchestration entity, it will not be a monolithic piece of functionality. Rather it will be realized as a collection of modular functions that integrates advances made in different domains like NFV, SDN or SON. Furthermore, it will use data-aided intelligence to optimize all aspects of service composition and delivery.

Network Slicing

A network slice, namely "5G slice", supports the communication service of a particular connection type with a specific way of handling the C- and U-plane for this service. To this end, a 5G slice is composed of a collection of 5G network functions and specific RAT settings that are combined together for the specific use case or business model. Thus, a 5G slice can span all domains of the network: software modules running on cloud nodes, specific configurations of the transport network supporting flexible location of functions, a dedicated radio configuration or even a specific RAT, as well as configuration of the 5G device. Not all slices contain the same functions, and some functions that today seem essential for a mobile network might even be missing in some of the slices. The intention of a 5G slice is to provide only the traffic treatment that is necessary for the use case, and avoid all other unnecessary functionality. The flexibility behind the slice concept is a key enabler to both expand existing businesses and create new businesses. Third-party entities can be given permission to control certain aspects of slicing via a suitable API, in order to provide tailored services.

Figure 9 illustrates an example of multiple 5G slices concurrently operated on the same infrastructure. For example, a 5G slice for typical smartphone use can be realized by setting fully-fledged functions distributed across the network. Security, reliability and latency will be critical for a 5G slice supporting automotive use case. For such a slice, all the necessary (and potentially dedicated) functions can be instantiated at the cloud edge node, including the necessary vertical application due to latency constraints. To allow on-boarding of such a vertical application on a cloud node, sufficient open interfaces should be defined. For a 5G slice supporting massive machine type devices (e.g., sensors), some basic C-plane functions can be configured, omitting e.g., any mobility functions, with contention-based resources for the access. There could be other dedicated slices operating in parallel, as well as a generic slice providing basic best-effort connectivity, to cope with unknown use cases and traffic. Irrespective of the slices to be supported by the network, the 5G network should contain functionality that ensures controlled and secure operation of the network end-to-end and at any circumstance.

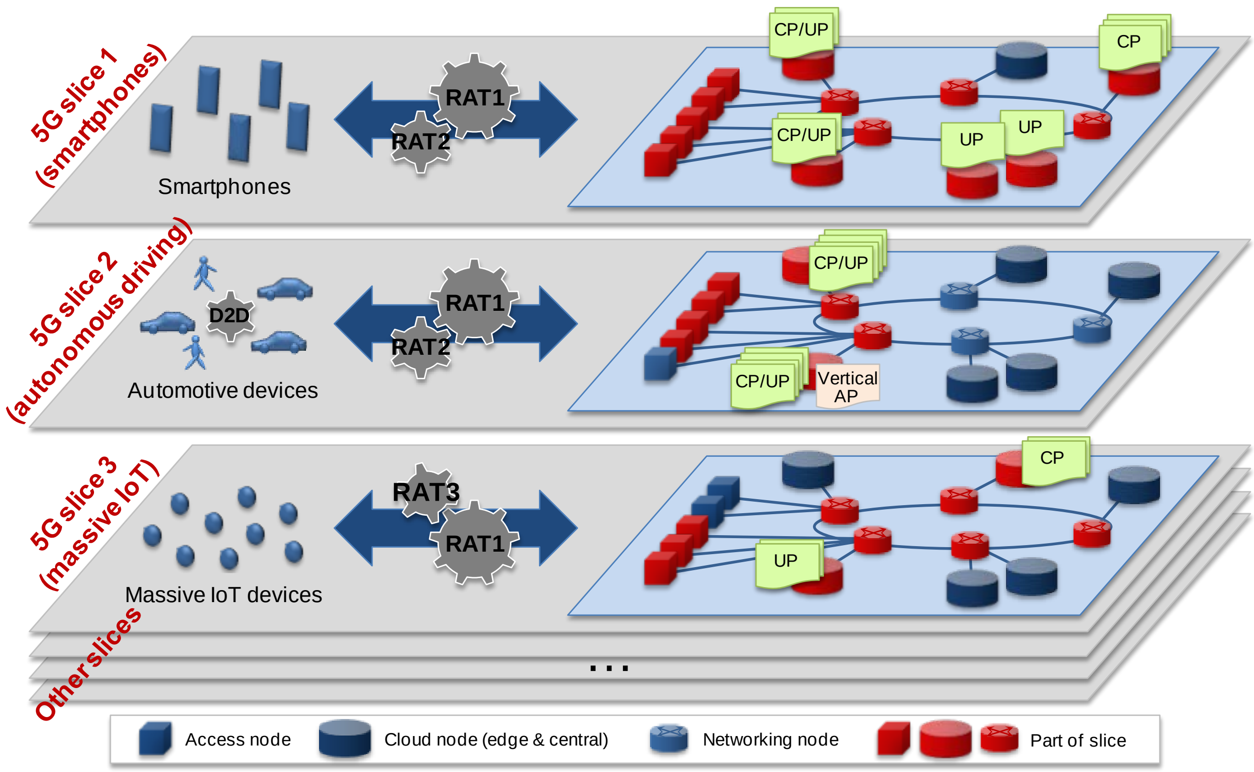


Figure 9: 5G network slices implemented on the same infrastructure

The use of both dedicated infrastructure resources for certain slices, as well as the use of shared infrastructure resources and functions between multiple slices is needed. One example of a shared function is the radio scheduler. The scheduler of a RAT will typically be shared among multiple slices, and will play a definitive role in allocating resources and setting the performance of a 5G slice, including the extent to which consistent user experience is realized. The scheduler implementation in today's networks is typically proprietary. Nevertheless, the required level of openness needs to be defined to have sufficient control over this critical function.

To realize such 5G system architecture, the C- and U-plane functions should be clearly separated, with open interfaces defined between them, in accordance with SDN principles. In addition, open interfaces should be defined between access-specific and access-agnostic functions so that additional access technologies, both fixed and radio, can be easily integrated into the 5G network in future. The fronthaul interface(s) between remote radio units and baseband units should also be open and flexible, offering multi-vendor operation and good forward and backward compatibility, while providing options for transport bandwidth reduction. Furthermore, the interfacing between functions must allow for multi-vendor provisioning of different functions.

An important consideration in such system architecture is the granularity at which functions are defined. While finer granularity will improve flexibility, it can also lead to significant complexities. The testing efforts for different function combinations and slice implementations will be cumbersome, and interworking issues among different networks will arise. Therefore, the right granularity to balance the goals of flexibility with complexity needs to be identified. This will also influence how the eco-system delivers solutions.

5G System Components

The architecture and principles described above lead to emergence of a set of key components and terminology of a 5G system, as described below.

5G RAT family (5GRF): As a part of the entire 5G system, the 5G RAT family is the set of one or more standardized 5G RATs that together support NGMN 5G requirements. The 5G RAT family should provide wide coverage, as this is a critical factor for marketing new technology.

5G RAT (5GR): A 5G RAT is a component radio interface of the 5G RAT family.

5G Network Function (5GF): A 5G network function (5GF) provides a particular capability to support communication through a 5G network. 5G network functions are typically virtualized, but some functions may be provided by the 5G infrastructure using more specialized hardware. The 5GFs comprise RAT-specific functions and access-agnostic functions, including functions to support fixed access. The 5GFs can be classified into mandatory and optional functions. Mandatory functions are common functions necessary for all use case categories, e.g., authentication and identity management. Optional functions are the functions that are not always applicable for all the use cases. For instance, a mobility function like handover may be used only for the mobile broadband use case category and not at all for the low-end machine communications category. Optional functions may also have different variants tailored to the traffic type and use case.

5G Infrastructure (5GI): The 5G infrastructure (5GI) is the hardware and software basis for the 5G network, including transport networks, computing resources, storage, RF units and cables supporting the network functions providing the 5G network capabilities. 5G RAT(s) and 5GFs are implemented or realized using the 5GI.

5G End-to-end Management and Orchestration Entity (5GMOE): The 5G end-to-end management and orchestration entity (5GMOE) creates and manages the 5G slices. It translates use cases and business models into concrete services and 5G slices, determines the relevant 5GFs, 5GRs and performance configurations, and maps them onto the 5GI. It also manages scaling of the capacity of individual 5GFs and their geographic distribution, as well as OSS and SON.

5G Network (5GN): A 5G network is the 5GFs, 5GRs, the associated 5GI (including any relaying devices) and the 5GMOE supporting communication to and from 5G devices. In other words, a 5G network is realized when a 5G RAT utilizes any subset of functions from the 5GFs implemented on the 5GI to support communications with a 5G device. On the contrary, the network created when the 5GFs are used to support communications with a 5G device through a non-5G RAT is not considered as a 5G network.

5G Device (5GD): A 5G device is the equipment used to connect to a 5G network to obtain a communication service. 5G devices can support machines as well as human users.

5G System (5GSYS): A 5G system is a communications system comprising a 5G network and 5G devices.

5G Slice (5GSL): A 5G slice is a set of 5GFs and associated device functions set up within the 5G system that is tailored to support the communication service to a particular type of user or service.

5.5 5G Technology Options

5G Radio Access Technology Options
Theoretically, two options can be foreseen for 5G RAT. In the first, a single unified RAT that can be optimized for different frequencies and use cases through parameter configuration of a common air interface can be designed. With this approach, a single RAT provides the minimum set of defined 5G features for different combinations of use case categories. This approach would be ideal for the long term, as operators could avoid management of multiple access networks. Device implementations are expected to be simpler, if a single unified RAT covers all the use cases and frequencies. However, it is foreseen to be technically very challenging to design an air interface with such degree of flexibility while maintaining performance and efficiency over the wide range of use case categories and frequency bands. Furthermore, introducing a unified RAT because of the need to support some new use case categories (e.g., low-end machine communications) may require re-farming of existing spectrum already used for legacy technologies such as LTE-Advanced.

Depending on the particular use case categories driving the need for such migration, the cost of migration to a single unified RAT needs to be carefully compared with the achievable benefits. Even if the new RAT improves the overall spectral efficiency, significant gains will not be realizable for the mobile broadband use case category by only re-farming portions of existing 4G spectrum (e.g., 20 MHz). In fact, the performance may very well be lower than that achieved with LTE-Advanced with carrier aggregation in such a scenario. This could detract investment in the new RAT, given that LTE-Advanced devices with higher capabilities to aggregate up to 100 MHz of spectrum may have high penetration in the network by 2020. Thus, aggregation with LTE/ LTE-Advanced (e.g., carrier aggregation and dual connectivity) will be fundamental to achieving good performance during the initial migration phase to the new RAT. Furthermore for operators without significant 4G investments, the thought of a new unified RAT with potentially better wide coverage compared to LTE/ LTE-Advanced may demotivate ongoing deployment of LTE/ LTE-Advanced, which may cause detrimental effects on roaming. Such implications of migration need to be carefully considered, to make introduction of the new RAT economically viable.

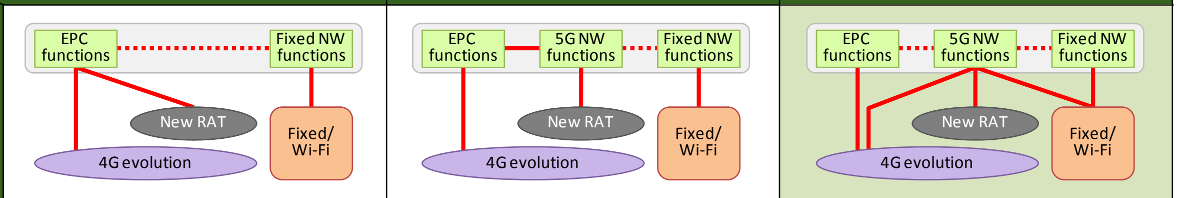
Alternatively, multiple RATs with potentially different air interfaces could complement each other, acting as a single unit. Compared to the previous alternative, this approach allows for the design and phased deployment of multiple use-case-category-specific or spectrum-specific RATs, which may be technically easier to achieve and more economically viable. A new RAT could be motivated by high carrier frequencies (e.g., bands above 6 GHz), lower latency, and specific use cases. For instance, a RAT optimized for use in higher frequency bands could be used to provide capacity and high data rates in dense urban areas, as well as indoors, to support the mobile broadband use case category. Due to the limited coverage of higher frequency bands, another RAT optimized for use in lower frequency bands, and suitable for multiple use case categories (for example, LTE evolution after a certain 3GPP Release, e.g., with features to efficiently support some types of machine communications) can be used in a complementary manner to provide coverage. As another example, a specialized RAT could be designed to support ultra-low latency or ultra-high reliability and can be used in combination with another RAT suitable, e.g., for mobile broadband use case to meet the requirements of a diverse set of 5G use case categories.

The exact combination of use case categories that will be deployed by an operator will depend significantly on the operator's business model. Hence, defining the minimum set of features that characterizes 5G RAT at this stage is difficult. Moreover, further evaluations on technologies need to be performed to better understand sensible clustering of use case categories and spectrum, to define the potential new RATs. Therefore, the term "5G RAT family" is defined to cover a set of one or more standardized 5G RATs that together support NGMN 5G requirements, and the industry is encouraged to study further to decide the 5G RAT(s) to be standardized.

Given the wide spectrum of capabilities that need to be delivered, 5G will most likely be comprised of more than one RAT, each optimized for certain use cases and/ or spectrum. However, new RATs should not be unnecessarily defined to achieve some niche optimization. The number of RATs should be minimized, ideally down to one, to achieve economies of scale. Even if multiple RATs are to be defined, commonality should be achieved to the largest extent possible. For example, the protocol stack above layer 2 should be harmonized, and the same control functionality should be applied, thereby making the different RATs rather different modes of operation of a single RAT.

5G Interfacing Options

Three interfacing options for the access technologies, as depicted in Figure 10, provide potential migration paths towards 5G. It is assumed here that the 5G RAT family comprises multiple RATs optimized for different use case categories and/ or spectrum. In Figure 10, the 5G RAT family comprises a new RAT (e.g., optimized to provide high data rates and capacity in higher frequency bands, ultra-low latency or ultra-high reliability, among others) and an evolved LTE RAT (e.g., after a certain 3GPP Release, to provide coverage and support for other use case categories such as low-end machine communications). Nevertheless, the discussion in this section also applies to the alternative case of a new unified RAT supporting all scenarios.



	Option 1	Option 2	Option 3
Pros	▪ No changes to 4G RAN ▪ No need for revolutionary 5G NW functions design	▪ No changes to 4G RAN ▪ 5G NW functions/ new RAT design can be optimized to fully benefit from new technologies (e.g., virtualization)	▪ 5G NW functions/ new RAT design can be optimized to fully benefit from new technologies (like virtualization) ▪ Solves mobility issues of option 2 ▪ Provides a sound migration path
Cons	▪ Tied to the legacy paradigm for all the use cases (which may be expensive)	▪ New design could only be utilized where there is new RAT coverage ▪ Potential signalling burden due to mobility if the new RAT does not provide seamless coverage	▪ Potential impact on legacy RAN to operate concurrently with legacy CN functions and 5G NW functions.

NW Network
EPC Evolved packet core
RAN Radio access network
RAT Radio access technology

Defined interface/ reference point (solid line)
Potential interface/ reference point (dotted line)

Figure 10: Access-technology interfacing options[2]

In the first interfacing option, all access-agnostic components supporting the 5G RAT family are provided through EPC[3]. This option may require evolution of EPC to enable 5G access-agnostic functions to be provided. With this option, there is minimal impact to legacy RAN. Nevertheless, the

[2] The diagram presumes some LTE functions supporting the interface to EPC are not 5GFs, but performed at the RAT level. The exact scope of the 5GFs in option 3 needs further consideration.
[3] In this case the RAT-specific functions are assumed to be implemented at the RAT level.

degrees of freedom to evolve the EPC in a manner that efficiently provides 5GFs to support the diversity of use cases may be limited. Thus, legacy paradigms may be applied to all use cases, which may be inefficient and expensive.

In the second option, the 5G access-agnostic functions are provided both through an evolution of EPC and a new design denoted “5G NW functions”. But the new design only supports the new RAT and 4G evolution is supported by the EPC. The advantage is that it allows the benefits of new technologies such as virtualization to be realized while at the same time minimizing the impact to legacy RAN. However, the drawback is that the benefits of the new design can only be realized in areas where there is new RAT coverage. Furthermore, due to limited coverage of the new RAT, interworking interfaces may be needed between the new design for 5GFs and EPC to support mobility between the new RAT and 4G evolution. Providing mobility support through such interfaces may incur significant signalling burden.

In the final option, all components of the 5G RAT family are supported by the new 5GFs design. Other RATs (e.g., Wi-Fi) and the fixed network may also be supported through the new 5GFs design. This option also allows for support of the 4G evolution through the EPC to provide backward compatibility for devices that cannot utilize the new design (e.g., devices that only support LTE before a certain 3GPP Release). Similar to Option 2, this option allows the benefits of new technologies to be fully realized. In addition, it overcomes the mobility issues associated with Option 2. This is because mobility between the new RAT and 4G evolution can be handled by the 5GFs without the need for any interworking. In addition, this option provides a sound migration path, since all RATs (4G evolution as well as evolution of local-area access technologies) can immediately benefit from the 5GFs, even in areas without new RAT coverage.

Nevertheless, Option 3 also introduces new challenges. For instance, it requires the 4G RAN to be upgraded to support both connection through the EPC and the new 5GFs design. Some interfaces may also be needed during the migration phase for basic limited interworking until all 4G base stations have been upgraded to support the new 5GFs design. The same is true for the fixed network and other RATs which will connect to the new 5GFs design. Nevertheless, supporting multiple-connectivity at the device side should reduce the legacy interworking requirements on the network side, and allow design of 5GFs without legacy constraint. For these reasons Option 3 is currently considered by NGMN as the preferred option. In order to facilitate migration toward 5G, NGMN recommends that LTE/ LTE-Advanced and Wi-Fi, as well as their evolution, are to be supported by the new 5GFs design. Thus, the access-agnostic network functions should accommodate any new RATs, as well as LTE/ LTE-Advanced, Wi-Fi, and their evolution.

Regardless of the architecture option pursued, harmonizing different identity and authentication paradigms in cellular networks, (wireless) local access networks, and fixed networks will be essential to enable convergence of different access types, and also to facilitate the realization of different business models. The architecture must also facilitate further convergence of fixed and mobile networks in a manner that efficiently addresses the needs and requirements originating from regulators (e.g., requirements to support MVNOs and MNP for different parts of the converged network).

6. SPECTRUM

6.1 Frequency Bands

6.1.1 Suitability of Existing Mobile Bands

NGMN's requirements for 5G include the need to support a wide range of applications that may have differing requirements for the underlying mobile connectivity. This will require access to a range of spectrum bands with differing characteristics in order to address a wide range of requirements for coverage, throughputs and latency in the most cost efficient manner and to make effective use of the spectrum.

5G will build on earlier generation mobile technologies and will bring additional capabilities. Spectrum bands already licensed to MNOs will form an essential foundation for 5G mobile services. It is therefore important to allow operators to "re-farm" existing spectrum bands to 5G technology according to their deployment strategy. This will enable improvements in spectrum efficiency to be achieved and new capabilities to be introduced. It will also enable the necessary long-term investments to be planned.

6.1.2 Wireless Spectrum Needs

6.1.2.1 Additional Network Spectrum Requirements

It is expected that 5G will be integrated under the umbrella of the International Mobile Telecommunication (IMT) family that is developed within the framework of the International Telecommunications Union (ITU). From the MNOs' perspective, predictable integration of new standards is of pivotal importance, ensuring a harmonized standardization process and a stable regulatory environment. Additional spectrum allocations to support 5G requirements should be identified within the global framework provided by the ITU Radio Regulations and implemented in regional and national allocation and assignment decisions.

It is anticipated that the ITU World Radiocommunication Conference 2015 (WRC-15) will identify new bands for IMT that will become available in addition to spectrum bands already in use by mobile networks or for which licences have recently been awarded. However, it is likely that still further spectrum will be needed to deliver all the services described in the 5G vision. Such spectrum should be harmonised as widely as possible to ensure that 5G systems have global scale, thereby avoiding technical complexity both in network and terminal side.

The potential further spectrum requirements should be considered at the World Radiocommunication Conference that follows WRC-15, which would need studies on the following topics:

- Spectrum requirements for mobile services in the period beyond that considered by WRC-15.
- The need to evaluate possible candidate bands in higher frequencies to address new spectrum beyond the year 2020 for ultra-dense networks. Such frequencies are needed to allow very wide bandwidth channels to support very high data rates and short-range mobile connectivity (e.g. 500 - 1000 MHz of contiguous spectrum per network to support the multitude of services described in section 3.2.1) Total spectrum requirements should take in to account the potential need to accommodate multiple networks Therefore it is proposed to study technical feasibility of the ranges between 6 GHz and around 100 GHz, in particular those where primary or co-primary allocation to mobile in the ITU Radio Regulations exists already. The lower limit for the band range (above 6 GHz) should be further assessed.

Depending on the outcome of WRC-15, there may also be a need for a future Conference to consider further spectrum for mobile broadband, both for coverage and capacity. Spectrum below 1GHz-is particularly useful for coverage especially indoor and in rural areas, while spectrum above 6GHz is particularly useful to support very high data rates and short-range connectivity.

Whilst in a 5G context access to additional spectrum above 6GHz is of interest, it should be emphasized that in general low frequency spectrum (below 6GHz), especially sub-1GHz, is absolutely essential for an economical delivery of mobile services and this holds true for existing systems as well as future 5G systems. Therefore, priority must be put on how to make more spectrum in those low bands available, and how to use that spectrum much more efficiently.

Work on IMT/5G spectrum must not affect ongoing process towards identification of additional spectrum for LTE evolutions by WRC-15.

6.1.2.2 Need for backhaul network spectrum

With the inclusion of higher bands to accommodate small cells and new deployment architectures, and likely densification of 5G networks vis-à-vis prior generation systems, there is also likely to be a need for consideration of associated backhaul needs. In addition to fixed line backhaul solutions, for some scenarios wireless backhaul solutions using in-band or out-of-band spectrum may be required. This is especially relevant where the commercial business case of cells is such as to preclude additional costs associated with cutting pavements and accessing/installing conduits for cable/fibre-based backhaul infrastructure.

6.2 Spectrum Management Options

6.2.1 Continuing Need for Licensed Spectrum

Spectrum should normally be made available on a national basis with exclusive licences assigned to network operators, to enable quality of service to be managed. This needs to be harmonised on a global basis to support global roaming and generate economies of scale. Exclusive licensing regimes should remain the main and preferred solution for accessing core spectrum. The way to create sustainable consumer benefits and increased competition should start by creating regulatory and legal certainty in the market.

6.2.2 Supplementary Spectrum for Flexibility and Capacity

For some applications, the core exclusive spectrum will need to be supplemented by access to additional spectrum on a shared basis in order to deliver extra capacity for the best possible user experience in a consistent manner and in line with what customers require. Thus, in addition to exclusive licensed spectrum, some shared spectrum may also be required, for example:

- Additional licensed spectrum made available by an incumbent governmental/public user within a defined area and/or for a defined time for use by mobile operators.
- Licence-exempt use of spectrum may be a useful supplement for certain applications.

Access technologies utilising spectrum on a licence-exempt basis may continue to play a role in the data traffic management (offload purposes in various hotspot locations). However, the technical limitations in service management, cell coverage, and traffic handling, make them unlikely as substitutes for 5G, but rather may be more seamlessly integrated into the overall 5G platform. The specific technologies that may be relevant for use on a licence-exempt basis in the 2020 timeframe may include evolutions of WiFi as well as other air interface technologies. It will therefore be necessary to consider how the various technologies will coexist and share access to spectrum when operating on a licence-exempt basis and to address this in standards.

Provided that exclusive licensing regimes should remain the main and preferred solution for accessing spectrum, NGMN sees all additional spectrum management options as potentially relevant to IMT/5G and recommends that they will be further studied in order to maximize the access to the spectral resources. Furthermore if shared spectrum can be accessed in a way that gives some assurance of quality this is of particular interest.

6.2.3 Benefits of Spectrum Flexibility

Any spectrum and spectral efficiency gains as a result of flexibility in spectrum usage should be explored. These include:

- Continuing to increase the exclusive licensed spectrum with emphasis on improving the harmonization of the spectrum in regional and global scale
- Explore flexible utilization of MNO's licensed bands:
    - Optimized coexistence with other radio technologies and dynamic use of radio resources by different radio access technologies according to the scenarios, traffic load, user requirements, coexistence environment and etc.
    - Smart carrier aggregation to benefit from any spare frequencies.
    - Spectrum trading between operators
- Managing access to supplementary spectrum on a licensed shared and on a licence-exempt basis
    - Coordination mechanisms to manage the equitable access to shared spectrum by MNOs to maintain the high spectrum efficiency and ensuring interference is adequately controlled and managed.

6.3 Required Next Steps on Spectrum

Spectrum management changes typically require long timescales and involve a wide range of international stakeholders. The NGMN Alliance will seek to contribute constructively to the work of the various global, regional, and national regulatory bodies, in charge with shaping the spectrum policies. The ITU work on IMT and its associated World Radiocommunication Conference preparation are an important example of external activity where NGMN Alliance will have an interest in ensuring that its operators' 5G requirements are considered.

7. IPR

7.1 Business Objectives

NGMN is developing recommendations and an implementation strategy supporting a more transparent and predictable IPR eco-system for 5G Standards Essential Patents (SEP) across industries that will support commercial and sustainable implementation of 5G technologies and ensure that innovation is stimulated and innovators appropriately rewarded.

One of the business objectives is to make 5G access affordable for all types of devices from the high-end smartphones and tablets down to the low-end MTC (Machine Type Communication) devices such as smoke detectors and sensors. In order to ensure this objective, the IP licensing terms and conditions for 5G market should be fair, reasonable and non-discriminatory so to ensure the sustainable and successful mass deployment of any 5G service or Product Type including MTC devices to support the Internet of Things (IoT).

7.2 Proposal

NGMN proposes the following recommendations to the IPR eco-system for 5G:

1. **Improve 5G Standard Essential Patent (SEP) Declarations**
   NGMN recommends improving the existing structure and framework across the industry for Standard Essential Patent declarations in order to improve transparency and limit abusive patent declarations related to 5G standards, while still encouraging early declarations.

2. **Establish Independent 5G Standard Essential Patent (SEP) Assessments**
   NGMN recommends submitting each patent considered as a 5G Standard Essential Patent to an independent essentiality assessment prior to licensing in order to ensure quality declarations.
   To ensure transparency and effectiveness, NGMN recommends that each patent holder share the result of these independent essentiality assessments.

3. **Explore and establish Patent Pool licensing for 5G**
   NGMN recommends exploring and establishing an appropriate 5G patent pool framework.
   NGMN recommends that the 5G SEP holders determine appropriate licensing terms and conditions (including royalties) within the 5G patent pool framework to meet the overall NGMN business objectives.

7.3 Next steps

From April 2015, NGMN will engage with all industry partners to develop implementation plans for each of the recommendations stated above. Each implementation plan is to be ultimately approved by the NGMN Board.

8. WAY FORWARD

8.1 Introduction

The publication and dissemination of the NGMN 5G White Paper is only the initial step on the roadmap towards 5G launch and deployment. The key objective of NGMN's future activities will be that the commercial 5G solutions will fulfil the NGMN requirements. In order to reach this objective, detailed milestones and a roadmap needs to be defined and also the necessary steps on the roadmap have to be outlined. Guiding principle should be that the migration towards 5G and the development of 5G solutions should be as efficient as possible for the operator and supplier industry stakeholders, and as beneficial and seamless as possible for the end-user.

The "Way Forward" section of the White Paper outlines the roadmap towards 5G development and rollout. Furthermore, it defines the necessary steps to be taken in order to reach the given milestones. The section also highlights the roles and tasks of NGMN as well as the roles and tasks of other industry stakeholders.

8.2 Roadmap

The roadmap, milestones and steps to be taken towards the final deployment are essential prerequisites for the overall success of 5G. NGMN has defined a 5G roadmap that shows an ambitious time-line with a launch of first commercial systems in 2020. At the same time it defines a reasonable period for all the industry players to carry out the required activities (such as standardisation, testing, trials) ensuring availability of mature technology solutions for the operators and attractive services for the customers at launch date. The key milestones are as follows:

- Commercial system ready in 2020
- Standards ready end of 2018
- Trials start in 2018
- Initial system design in 2017
- Detailed requirements ready end of 2015

The launch of 5G will happen on an operator and country specific basis. Some operators might plan to launch in 2020 – others will plan for a later deployment. The roadmap represents the baseline planning from an NGMN perspective and milestones might be shifted in the course of the 5G development due to external factors (e.g. standardisation process, etc.). The detailed roadmap is shown below:

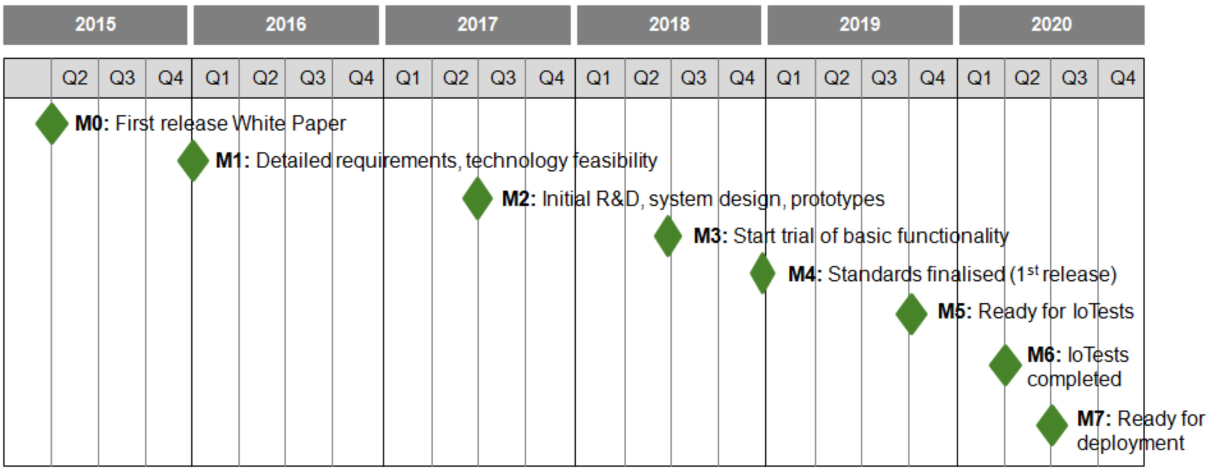


Figure 11: NGMN 5G Roadmap

M0: First release (v1.0) NGMN 5G White Paper, NGMN Industry Conference: Vision, use cases, requirements, architecture, spectrum, IPR
M1: Detailed requirements available, technology feasibility and options explored
M2: Initial R&D and system design done, first prototypes ready, study and recommendation for standardisation available, standardisation starts
M3: Trial of basic functionality starts
M4: Standards finalised (1st release)
M5: Infrastructure and terminals ready for interoperability tests and certification. Start of friendly customer trials
M6: Infrastructure and terminals interoperability tests completed
M7: First commercial infrastructure. Services and terminals ready for deployment

8.3 Main Industry Stakeholders, Roles and Activities

The 5G communications system comprises the whole end-to-end system: A 5G network and the associated user equipment. Numerous groups of industry stakeholders like vendors, researchers, standards developing organizations (SDOs), certification bodies and others are involved in the development of 5G. It will be essential for successful development and launch of 5G that all relevant industry stakeholders are aligned and that guidance is received from NGMN on the NGMN 5G requirements and roadmap. NGMN will continuously monitor and support the 5G related stakeholder activities and will provide the necessary input (for details on the NGMN role and activities, see section 8.4)

This sub-section describes the expected role and activities of the main 5G industry stakeholders. Furthermore, it outlines the requirements and guidelines from an NGMN perspective for the different activity areas.

8.3.1 Standardization

NGMN sees the standardization of technology as essential for the global success of the future 5G solutions and the related ecosystem. Standardization ensures (multi-vendor) interoperability and economies of scale. Furthermore, it minimizes the complexity and thereby reduces the cost of interfaces.

Given the range of interfaces, network elements and legacy systems, numerous standardization bodies (with many contributors) are expected to get involved in the 5G standardization work. There is the risk of conflicting standards, redundant options and development delays due to the diversity of interests of the involved parties. It needs to be ensured that

- Parallel work on similar areas (with potentially conflicting standards) is avoided.
- Solutions of different organisations are harmonized.
- Options are reduced to an absolute minimum set.

SDOs should develop standards with the ultimate aim to support the timely delivery of competitive products, which will meet the needs of mobile operators and their customers.

NGMN encourages the industry to strongly contribute to the 5G standardization work to ensure the timely and successful development of 5G solutions.

8.3.2 Certification

NGMN sees the certification of terminals within a 5G terminal certification regime as essential to provide the assurance that terminals will perform correctly on networks when being launched. Close co-operation of all related industry stakeholders (SDOs, certification bodies, test industry, terminal manufacturers) is needed to ensure timely availability of certified terminals.

It is recommended that test specifications, test equipment and test cases are developed as much as possible in parallel to the core specifications. Test specifications should be released as close as possible to the release of respective core specification.

8.3.3 Research and Development

Numerous global vendors, research institutes, industry initiatives, and NGMN advisors are involved in the research and development of 5G solutions up to 2020 and beyond. They will serve different markets and customers with different needs (in terms of time-lines, features etc.) and legacy.

Close alignment and development based on NGMN requirements is needed to ensure the successful and timely deployment and launch of 5G.

Open source initiatives that ensure reliability, quality and interoperability are also to be qualified to speed up the development of 5G.

8.3.4 Other Activity Areas

There are other prerequisites that need to be fulfilled to ensure the success of 5G:

- Roaming and interconnect of networks, terminals
- Sufficient, global spectrum
- Integration / APIs for services
- and others

There are well established international industry organisations with specific capabilities enabling them to effectively address those items. Close alignment of the related tasks is needed and it will be essential that the development of deliverables will happen in line with NGMN requirements.

8.4 NGMN Role and Activities

In its initial phase, the NGMN Alliance successfully enabled the launch of commercial LTE services in 2010 through its activities on technology, spectrum, IPR, ecosystem and trials. To enable the launch of commercial 5G services as outlined in the roadmap, NGMN will build on its unique strengths and characteristics:

- Positioning: Well established and growing partnership of worldwide leading operators and vendors – strong, international industry backing
- Role:
    - Open perspective – activities across standards, co-operations with international fora
    - Business driven – consideration of end user demand and industry needs
    - Operator requirements – requirements defined by operators, solutions identified by vendors, research perspective delivered by academia
- Scope:
- Clear focus on next generation technology – covering end-to-end system and services aspects
- Working on requirements level – standards being developed by SDOs
- Organisation: Decision oriented governance, lean organisational structure, project oriented approach, efficient processes
- Working procedures: Results based on NGMN Partner contributions – projects under the lead of NGMN Partner representatives and staffed with experts from Partner companies

This sub-section of the "Way Forward" section outlines the main tasks and activities to be carried out by NGMN in order to enable the successful 5G launch based on the NGMN requirements and the given roadmap. NGMN will define a work-programme covering those tasks and activities – the initial 5G work-programme will be set-up immediately after publication of the NGMN White Paper.

8.4.1 Definition of Requirements and Performance Targets

Main task within the NGMN work-programme is the development of end-to-end requirements for next generation technology from an operator perspective and with a clear business view.

In the White Paper, NGMN outlined initial qualitative and quantitative 5G requirements in the areas of user experience, system performance, enhanced services, business models and management & operations.

Going forward, the following tasks will be carried out by NGMN:

- Detailing and enhancement of existing White Paper requirements, identification of potential gaps (requirement areas not addressed so far) and definition of new requirements. Consolidation, prioritisation and documentation of those requirements
- Close co-operation and liaising with relevant industry organisations – distribution, dissemination of requirements to ensure acknowledgement and implementation. In addition, public communication of results and communication within NGMN.

8.4.2 Analysis of Technical Solutions, Assessment of Feasibility, Alignment and Guidance on Critical Issues

NGMN will closely work together with all its Partners (Members, Sponsors, and Advisors) to analyse the available technical 5G solutions in terms of performance and capabilities. It should be assessed whether (and to what extent) NGMN requirements will be met and where the major challenges are. Furthermore, it should be evaluated whether there are any diverting views or options that will lead to challenges and increased complexity in the development and implementation. In addition, it should be checked how potential 5G solutions relate to NGMN's design principles and architecture options.

This task in the NGMN work-programme will be done based on regular workshops and project-based activities that will allow the close and continuous co-operation and exchange of information with all NGMN Partners.

Based on this analysis, NGMN will decide on appropriate measures and messages and it will give guidance to the industry – objective will be to align global industry solutions and to ensure development and implementation of solutions meeting the NGMN requirements.

8.4.3 Tracking and Driving of Standardization and Certification, Guidance on Critical Issues

Based on the NGMN requirements and guidance, industry standards should be developed defining solutions able to meet the NGMN requirements and to meet the NGMN roadmap milestones. NGMN will continuously monitor and review the status, time-line and content of the NGMN requirement related items in the standards. In case of any issues – performance gaps vs. requirements, diverging architecture options, delay in standards availability vs. NGMN roadmap etc. – NGMN will decide on appropriate measures and messages and it will give guidance to the industry on how to address the identified issues. It will be important that NGMN highlights standardization and certification priorities to the SDOs, certification bodies and other relevant industry stakeholders so that they could focus their work and will be able to meet the NGMN roadmap milestones.

8.4.4 Development of Use-Cases, Implementation Guidelines

The NGMN work is not limited to the documentation and dissemination of its views (requirements, architecture vision etc.). NGMN wants to make sure that the initial ideas will later-on be implemented in vendor solutions and will be deployed in the live-networks and terminals.

Based on the past experiences with the launch of new network generations and the expected future network deployment behaviour, NGMN will derive use-case and implementation guidelines for 5G technology. In this context, a "use-case" should be understood as a practical deployment scenario (deployment use case) for 5G technology (e.g. backhaul architecture options, radio access topologies).

Those NGMN use-cases will provide opportunities for vendors to learn about the operator expectations and priorities and to check this against the potential performance and behaviour of the solutions available. Together with specific NGMN implementation guidelines – jointly developed by NGMN Members, Sponsors and Advisors – those use-cases will support further enhancement of 5G solutions and will ensure the smooth and successful rollout of 5G solutions.

As in the past, NGMN will closely align its activities in the launch and implementation phase with all other relevant industry organisations in order to ensure alignment and to avoid overlap or duplication of work.

8.4.5 Sharing of Experiences, Analysis of Trial Results

NGMN Partners (Members, Sponsors, and Advisors) will carry out trials and testing of 5G solutions as soon as first prototypes are available (proof of concept) and standardization has started (interoperability and later friendly customer trials).

NGMN will accompany and monitor industry trial and testing activities and will assess the results of those efforts. Objective will be to evaluate whether the White Paper performance targets will be met by certain network functions or solutions. Furthermore, this activity will provide an early identification of potential performance issues and will provide the opportunity to give feedback on these findings to relevant industry stakeholders (vendors, standardization organizations …).

Besides this trial monitoring activity, NGMN will serve as platform for its Partners to continuously share experiences and information on testing, trialling and deployment (by building on the NGMN communication project infrastructure, workshops etc.).

8.4.6 Interaction of NGMN with Industry Stakeholders

To ensure the success of its work, NGMN in the past has already established co-operations with all leading industry organisations (e.g. ETSI, GSMA, TMF). This close collaboration will now be continued based on regular communication, monitoring and potential joint projects. NGMN will establish a separate work-item dedicated to the co-operation with the most relevant industry organisations.

Communication and co-ordination with other, non-NGMN Partners (such as regulators, governments) will happen based on liaison statements or based on the NGMN information that is publicly available.

Within the NGMN Alliance, i.e. together with its NGMN Partners (Members, Sponsors, and Advisors), NGMN will set-up projects, workshops and other proven ways of interactions for 5G. In parallel, NGMN will present its public deliverables at conferences and will distribute them via mailings and other communication channels.

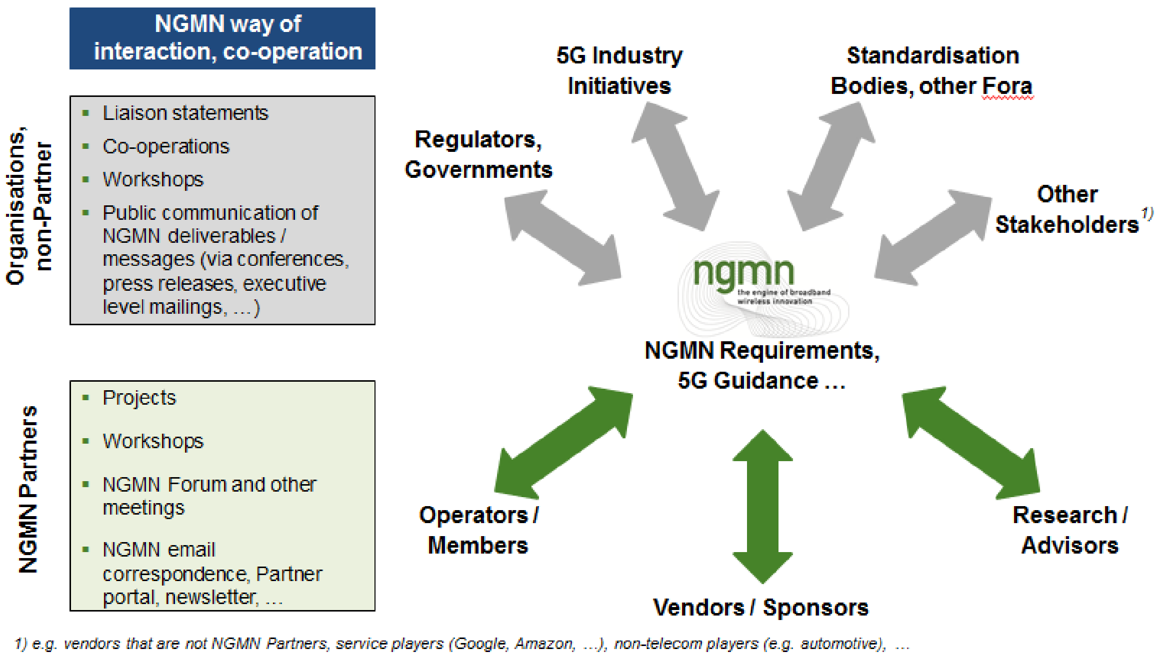


Figure 12: Interaction with Industry Stakeholders

9. CONCLUSIONS

5G is expected to have countless use cases, many unimagined today. In NGMN's vision, 5G is an end-to-end ecosystem to enable a fully mobile and connected society. It empowers value creation towards customers and partners, through existing and emerging use cases, delivered with consistent experience, and enabled by sustainable business models. The 5G vision has provided guidance to the definition of requirements, architecture and other aspects such as spectrum and IPR.

The 5G requirements cover the end-to-end considerations including user, system, enhanced service, management and operation, device and business model requirements. It is particularly important to keep improvements in the following areas in focus:

- **Network capability:** To cope with the diversity of use cases, the capabilities of the network need to be expanded to support e.g., higher data rates (>10x on average, >100x at cell edge), lower latency (>10x improvement) and higher connection density (>100x improvement). Nevertheless, not all capabilities need to be supported at the same time for the same user/use case. Thus, a flexible and scalable system that can steer those capabilities on demand is necessary.
- **Consistent customer experience:** Customer experience in 5G is defined by a set of customer-perceived and service-dependent experience metrics delivered consistently across time and service footprint.
- **Flexibility:** Supporting a wide range of use cases and business models requires 5G to provide a high degree of flexibility by design, along with trust, reliability and security. This applies for the level of modularity of the system as well as the granularity level for scaling the system on demand and as per need. Network resources and capability will be provided and allocated dynamically, on demand, per context, and in near real-time.
- **Efficiency:** 5G should show foundational shifts in cost and energy efficiency, and device power consumption, while supporting the expected traffic growth and the use cases. Sustainability and efficiency in deployments and management of potentially ultra-dense and multi-layer deployments is fundamental to the 5G eco-system.
- **Innovation:** The 5G eco-system is an open eco-system that enables innovations at a fast pace, involving many partners. 5G should provide the capabilities to allow this, with value creation for the operators and the market as a whole. Programmability of the network, availability of 5G value enabling capabilities (e.g. location, QoS, identity, security) and the related APIs are needed to make this happen.

5G should leverage an eco-system that is truly global, free of fragmentation and open for innovations. To this end, NGMN calls the industry to develop a single 5G standard based on the principles of open and global standards, utilising open interfaces and delivered based on globally available and non-proprietary solutions. Furthermore, The industry should ensure that roaming and interconnect in 5G are developed in a way that allows faster and more cost-efficient implementation, to ensure global reach of 5G and 5G services.

NGMN expects a more transparent and predictable IPR eco-system for 5G, to foster innovation, enable new use cases and business opportunities, and at the same time, reward innovators appropriately.

The commercial introduction of 5G will vary from operator to operator; however, NGMN encourages the ecosystem players to work towards a plan that would deliver globally and commercially available solutions by 2020.

LIST OF ABBREVIATIONS

3GPP	3G Partnership Project
5GD	5G Device
5GF	5G Function
5GI	5G Interface
5GMOE	5G Management & Orchestration Entity
5GN	5G Network
5GR	5G RAT
5GRF	5G RAT Family
5GSL	5G Slice
5GSYS	5G System
AAA	Authentication, Authorization, Accounting
ANDSF	Access Network Discovery and Selection Function
AP	Access Point
API	Application Programming Interface
APN	Access Point Name
ARPU	Average Revenue Per User
B2B	Business to Business
B2B2C	Business to Business to Consumer
B2C	Business To Consumer
BTS	Base Transceiver Station
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CN	Core Network
COMP	Coordinated Multipoint Transmission and Reception
CP	Control Plane
CPE	Customer Premises Equipment
C-PLANE	Control Plane
D2D	Device to Device
DAS	Distributed Antenna System
DL	Downlink
E2E	End to End
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
FDD	Frequency Division Duplex
FMC	Fixed Mobile Convergence
GSMA	GSM Association
HTC	Human Type Communication
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
HW	Hardware
IaaS	Infrastructure as a Service
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunications
IoT	Internet of Things
IPR	Intellectual Property Rights
ITU	International Telecommunication Union
KPI	Key Performance Indicator

LTE	Long Term Evolution
MIMO	Multiple Input Multiple Output
MME	Mobile Management Entity
MmW	Millimetre Wave
MNO	Mobile Network Operator
MNP	Mobile Number Portability
MTC	Machine Type Communication
MVNO	Mobile Virtual Network Operator
NaaS	Network as a Service
NFV	Network Function Virtualization
NGMN	Next Generation Mobile Network
NOMA	Non-Orthogonal Multiple Access
OPEX	Operational Expenditure
OSS	Operations Support Systems
OTT	Over The Top
PaaS	Platform as a Service
PCRF	Policy and Charging Rules Functions
P-GW	PDN-Gateway
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RAT	Radio Access Technology
RRU	Remote Radio Unit
SDN	Software Defined Networking
SDO	Standards Development Organization
S-GW	Serving Gateway
SINR	Signal to Interference plus Noise Ratio
SOA	Service Oriented Architecture
SON	Self-Organising Network
SSL	Secure Sockets Layer
SW	Software
TCO	Total Cost Of Ownership
TDD	Time Division Duplex
TMF	Telecom Management Forum
UE	User Equipment
UICC	Universal Integrated Circuit Card
UL	UL
U-PLANE	User Plane
VPN	Virtual Private Network
WRC	World Radiocommunication Conference
XaaS	Anything as a Service

ANNEX

ANNEX A: REQUIREMENT PER USE CASE CATEGORY AND NOTES

No. 1	Broadband access in dense areas	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 300 Mbps UL: 50 Mbps	This data rate is motivated by ubiquitous support of Cloud services, video and other digital services, possibly combined
E2E latency	10ms	
Mobility	On demand, 0-100 km/h	
Device autonomy	>3 days	
Connection Density	200-2500 / km^2	Total device density is 2000~25,000 / km^2, a 10% activity factor is assumed
Traffic Density	DL: 750Gbps / km^2 UL: 125Gbps / km^2	Connection density x User experienced data rate

No. 2	Indoor Ultra-high Broadband Access	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL 1Gbps, UL: 500Mbps	These data rates correspond to the Cloud storage service, which is the service with the highest data rates in the considered service mix (see the Traffic Density notes below).
E2E latency	10ms	
Mobility	Pedestrian	
Device autonomy	>3 days	
Connection Density	75,000/ km^2 (75 / 1000 m^2)	1 person per 4 m2, 30% activity factor; typical area is 500~1000m^2
Traffic Density	DL: 15 Tbps/km2 (15 Gbps / 1000 m2) UL: 2 Tbps/km2 (2 Gbps / 1000 m2)	A mix of services is considered: 25% of active users use Cloud storage services with data rates DL: 1Gbps, UL: 500Mbps 30% of active users use Desk cloud services with data rates DL: 20Mbps, UL 20Mbps 5% of active users use Multiparty video conferencing with data rates DL: 60Mpbs, UL: 15Mbps The remaining 40% of active users use less demanding services, which are neglected here. Within each service, an assumption is made on how much time is DL or UL. For cloud

		storage, DL is 4/5 of time while UL is 1/5; for desktop cloud the ratio for DL and UL are 5/6 and 1/6; while for the multiparty video conference the ratio for DL and UL are 1 and 1 (always on during an active video conference), respectively.

No. 3	Broadband Access in a Crowd	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 25Mbps UL: 50Mbps	Main use case is HD video/photo sharing
E2E latency	10ms	
Mobility	Pedestrian	
Device autonomy	>3 days	
Connection Density	150,000 / km^2 (30,000 / stadium)	Stadium: typical area 0.2 km2, 100000 persons, 30% activity factor
Traffic Density	DL: 3.75 Tbps / km^2 (0.75 Tbps/stadium) UL: 7.5 Tbps / km^2 (1.5 Tbps/stadium)	Connection density x User experienced data rate

No. 4	50+ Mbps everywhere	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 50Mbps UL: 25Mbps	Comfortable data rate allowing e.g. high resolution video combined with other digital services
E2E latency	10 ms	
Mobility	0-120 km/h	
Device autonomy	>3 days	
Connection Density	400 / km2 in suburban 100 / km2 in rural	
Traffic Density	DL: 20 Gbps / km2 in suburban UL: 10 Gbps / km2 in suburban DL: 5 Gbps / km2 in rural UL: 2.5 Gbps / km2 in rural	Connection density x User experienced data rate

No. 5	Ultra-low cost broadband access for low ARPU areas	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 10 Mbps UL: 10 Mbps	Broadband data rate allowing video, email and Web surfing
E2E latency	50 ms	
Mobility	On demand, 0-50 km/h	
Device autonomy	> 3 days	
Connection Density	16 / km^2	400 persons in a 25 km^2 isolated village, 10% activity factor

Traffic Density	DL: 16 Mbps / km^2 UL: 16 Mbps / km^2	Connection density x User experienced data rate DL: 400 Mbps / 25 km^2 UL: 400 Mbps / 25 km^2

No. 6	Mobile broadband in vehicles (cars, trains)	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 50Mbps UL: 25Mbps	
E2E latency	10ms	
Mobility	On demand, up to 500km/h	
Device autonomy	>3 days	
Connection Density	2000 / km^2 (500 active users per train x 4 trains, or 1 active user per car x 2000 cars)	Trains assumptions: 1000 persons per train; 50% activity factor; 2 trains per route (in opposite directions) within 1 km^2; 2 routes within 1 km^2; Cars assumptions (traffic jam case): 1000 cars are distributed over a 4-way x 4-way highway segment of 1 km length; 2 highways within 1 km^2; 2 persons per car; 50% activity factor.
Traffic Density	DL: 100 Gbps/km^2 (25 Gbps per train, 50 Mbps per car) UL: 50 Gbps/km^2 (12.5 Gbps per train, 25 Mbps per car)	Connection density x User experienced data rate

No. 7	Airplanes connectivity	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 15 Mbps UL: 7.5 Mbps	
E2E latency	10ms	
Mobility	Up to 1000 km/h	
Device autonomy	N/A	
Connection Density	80 per plane 60 airplanes / 18000 km^2	Assumption is a scenario with 20 planes in each of 3 sectors of the ground space
Traffic Density	DL: 1.2 Gbps/plane UL: 600 Mbps/plane	400 users per plane, 20% activity factor, leading to 80 active users per plane

No. 8	Massive low-cost/long-range/low-power MTC	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	Low (typically 1-100kpbs)	
E2E latency	Seconds to hours	
Mobility	On demand, 0-500km/h	Depends on use case
Device autonomy	Up to 15 years	Depends on use case
Connection Density	Up to 200,000/km^2	2 sensors per m^2 10% activity factor
Traffic Density	Not critical	

No. 9	Broadband MTC	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	See the requirements for the Broadband access in dense areas and 50+Mbps everywhere categories	
E2E latency		
Mobility		
Device autonomy		
Connection Density		
Traffic Density		

No. 10	Ultra-low latency	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 50 Mbps UL: 25 Mbps	
E2E latency	<1ms	
Mobility	Pedestrian	
Device autonomy	>3 days	
Connection Density	Not critical	
Traffic Density	Potentially high	

No. 11	Resilience and traffic surge	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 0.1-1Mbps UL: 0.1-1Mbps	Data rate for text, voice, or video messages
E2E latency	Regular communications: not critical; EWTS/PWS delivery time < 4s	
Mobility	0-120km/h	
Device autonomy	> 2 weeks	
Connection Density	10,000 / km2	Population density can go up to 80000 /km^2 (Daytime of Chiyoda ward, Tokyo), 15000 /km^2 (Average of Tokyo 23 wards). In case of disaster, a large number of

		persons will try to contact their relatives
Traffic Density	Potentially high	

No. 12	Ultra-high reliability & Ultra-low latency	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: From 50kbps to 10Mpbs UL: From a few bps to 10Mpbs	Allow some video transmission
E2E latency	1ms	
Mobility	On demand, 0-500km/h	Support of vehicular speeds is required for road safety applications
Device autonomy	Not critical	
Connection Density	Not critical	
Traffic Density	Potentially high	

No. 13	Ultra-high availability and reliability	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: 10Mbps UL: 10Mpbs	Data rate enabling real-time video and data transfers (e.g. maps)
E2E latency	10 ms	
Mobility	On demand, 0-500km/h	
Device autonomy	>3 days (standard) Up to several years for some critical MTC services	
Connection Density	Not critical	
Traffic Density	Potentially high	

No. 14	Broadcast like services	
Main Attributes	Requirement KPI	Notes
User Experienced Data Rate (also at the cell edge)	DL: up to 200Mpbs UL: Modest (e.g. 500kbps)	The maximum data rate can be used e.g. to distribute quickly 4K/8K movies, then cached at the device. Other broadcast like services can require a much lower data rate.
E2E latency	< 100ms	
Mobility	On demand, 0-500km/h	
Device autonomy	From days to years	Depends on the use case. MTC devices can need several years of autonomy
Connection Density	Not relevant	
Traffic Density	Not relevant	

ANNEX B: TECHNOLOGY GAP ANALYSIS TABLE

	No.	Requirement	Baseline (R12)	Target	Comment
User experience related requirements	1.2.1	Consistent user experience	No specific requirement defined	User experience should be virtually consistent across time, space and services. For HTC a minimum bit rate should be ensured For MTC consistency is not critical, minimum performance is more important and especially coverage and battery life	A suitable metric for consistency reflecting Quality of User. Consistent user experience is expected within specific use cases, but it will differ between the use cases. Experience and application specific requirements need to be defined.
	1.2.2	Data rate	Up to 100 Mbps average per user; peak of 600 Mbps (Cat 11/12 devices) and up to 4 Gbps (Cat 13 device) in low mobility (100 Mbps for high mobility)	Wide range, from a few bps for the Internet of Things (IoT) up to 1 Gbps DL / 500 Mbps UL sustained data rate	Depending on the available bandwidth, baseline performance figures for data rate range from 15 to 100 Mbps user data rates[4], with peaks of 1 Gbps for low mobility applications and 100 Mbps for high mobility. 5G data rate requirements are in the order of 100 Mbps to 1 Gbps for user experienced data rate, with peaks of tens of Gbps data rate. It is apparent that an improvement factor of approximately 10x can be expected for both average and peak data rates. It should be noted that the baseline LTE-A technology allows for much higher peaks, of 4 Gbps with 256 QAM modulation and 100 MHz worth of spectrum, but that is far from practical systems and the above figure

[4] Average user data rates can be derived by following two approaches. The first one comes from the spectral efficiency requirement given in 3GPP TR 36.912 for DL Urban macro (FDD) case (2.6 bps/Hz), and typically allocated user bandwidths between 5 MHz and 40 MHz. The second one comes from an empirical rule of thumb observed in practical deployments: typical peak to average user data rates stay in the range 4 to 6, therefore Cat 11/12 devices with 600 Mbps peak rate will enjoy average data rates in the order of 100 Mbps. Lower user bandwidths (e.g., 5 or 10 MHz) will still have typical user data rates in the order of up to 15 Mbps.

					has been provided taking into account more realistic deployment conditions.
	1.2.3	Latency	10 ms two-way RAN latency 50 ms end-to-end latency (content not at the edge) 25 ms end-to-end latency (content at the edge)	10 ms end-to-end user plane latency, and ~ 1 ms for ultra-high reliability, ultra-low latency services	Baseline latency of the RAN for the U-plane for a scheduled UE, assuming 0% HARQ BLER. Only some of the services will require a very low latency. For other services forcing a low latency could have a negative impact on other important requirements
	1.2.4	Mobility	High Mobility (up to 350 km/h for railway): 100 Mbps DL avg. data rate , 70 Mbps UL avg. data rate	Support of "user perceived infinite" Mobility-on-Demand based on use case from static (e.g. fixed sensor) to very high velocity (e.g. fast train), with at least 500 km/h to be supported. For airplanes, velocity of up to 1000 km/h will have to be supported	Baseline values taken from ITU requirements for IMT-Advanced in 3GPP TR 36.912, Tables 16.4.1.4-1 and 16.4.1.4-3.
	1.2.5	Device power efficiency	Not specified, left to implementation	Support of "user perceived no-charging". Significant improvement in power efficiency (use case specific): ▪ > 3 days autonomy for a smartphone] ▪ up to 15 years terminal autonomy for a low-cost MTC device ▪	Radio power consumption should be kept at certain level (e.g., 200 mA at full power full duplex).
System Performance Requirements	1.3.1	Coverage capability	▪ 1 ~ 3 Mbps at cell edge ▪ 100 km cell radius	Cost efficient throughput coverage, services available anywhere and anytime (Linked to consistent experience): ▪ Cell radius of [x] km ▪ Data rate [TBD] at a given coupling loss [TBD]	Note: Coverage requirements are already covered in "Consistent/Homogenous user experience" section and "Connectivity transparency" section – Suggestion to remove or integrate as may be applicable
	1.3.2	Connection density	Not explicitly specified, but typically ~2000 active users per	Support high density of connections where needed (i.e. X connections	Connection density is defined as the total number of connected devices per

			km2 can be supported in dense urban areas and/or stadiums	per given area): ▪ Up to 200,000 active connections per Km2 for Massive low-cost/long-range/low-power MTC use case category	unit area (exchanging data with the network). Note: in the use cases, the indicated figure does not include the background traffic. The requirements may in some cases be the sum of the requirements of the different use cases (e.g. HTC+MTC)
	1.3.3	Traffic volume density/Area capacity	Not specified	Should be able to handle very large traffic volumes and variations (Should be able to scale independently of connection density): ▪ Up to 15Tbps/km2 DL and 2Tbps/Km2 UL for Indoor ultra-high broadband access use case category	The requirements may in some cases be the sum of the requirements of the different use cases (e.g. HTC+MTC)
	1.3.4	Spectrum efficiency	Baseline spectrum efficiencies are use case dependent: Indoor (FDD): DL: Cell average: 6.1 b/s/Hz, cell edge 0.24 b/s/Hz UL: Cell average: 4.3 b/s/Hz, cell edge 0.25 b/s/Hz Indoor (TDD) DL: Cell average: 6.1 b/s/Hz, cell edge 0.22 b/s/Hz UL: Cell average 3.9 b/s/Hz, cell edge 0.25 b/s/Hz Microcellular (FDD) DL: Cell average: 3.1 b/s/s/Hz, cell edge 0.11 b/s/Hz UL: Cell average 2.5 b/s/Hz, cell edge 0.09 b/s/Hz	Spectrum efficiency should be significantly higher than in 4G, with the emphasis on average and cell-edge spectrum efficiency (linked to throughput & capacity): ▪ The average and cell edge spectrum efficiency should be an order of magnitude higher than those targeted for 4G networks ▪ Ultra-high data rate use case (e.g. indoor or hotspot) ▪ High data rate use case (e.g. indoor or outdoor macro cell scenario) ▪ Medium to low data rate use case (e.g. wide area coverage) ▪ Machine-to-machine use cases (e.g. Internet of Things, meters, …)	Baseline spectrum efficiency depends on the use case as presented in TR 36.912. KPI is the data throughput per unit of spectrum resource per site. There are 5G critical communications and user safety use cases with high coverage and service reliability requirements. For them, the spectrum efficiency may be relaxed

			Microcellular (TDD) DL: Cell average: 4.2 b/s/s/Hz, cell edge 0.09 b/s/Hz UL: Cell average 2.8 b/s/Hz, cell edge 0.07 b/s/Hz Urban macro (FDD) DL: Cell average: 2.6 b/s/s/Hz, cell edge 0.074 b/s/Hz UL: Cell average: 2.1 b/s/Hz, cell edge 0.1 b/s/Hz Urban macro (TDD) DL: Cell average: 2.6 b/s/s/Hz, cell edge 0.075 b/s/Hz UL: Cell average: 2.0 b/s/Hz, cell edge 0.1 b/s/Hz High-speed, Rural macro (FDD) DL: Cell average: 3.5 b/s/Hz, cell edge 0.1 b/s/Hz UL: Cell average: Cell average 2.3 b/s/Hz, cell edge 0.13 b/s/Hz High-speed Rural macro (TDD) DL: Cell average: 3.2 b/s/Hz, cell edge 0.09 b/s/Hz UL: Cell average: 2.5 b/s/Hz, cell edge 0.15 b/s/Hz		
	1.3.5	Resource and Signaling efficiency	No specific requirements for applications other than MTC. Baseline network capabilities for MTC include: ▪ To reduce peak data/signaling traffic from very large numbers of MTC	Signaling should not represent a higher volume than data, and the associated overhead should be justified by the application needs: ▪ Max reusability, portability and sharing of functions and resources	Sporadic, small packets and massive connections result in high signaling overhead. 5G should significantly reduce signaling. Possible KPI: ratio of signaling to data payload. Baseline taken from 3GPP TS 22.368.

			devices ▪ To keep connectivity for large numbers of MTC devices ▪ To lower power consumption of MTC devices ▪ To reduce the frequency of mobility management procedures and location updates per MTC device ▪ To allow/reject MTC access requests outside a defined access grant time interval, or after a defined access duration ▪ To defer access by delay-tolerant MTC devices in case of congestion	▪ Resource & traffic optimization ▪ Significant signaling reduction For certain MTC use cases (e.g. stationary devices), not all control plane functionalities and related signaling may be needed. Managing/avoiding signaling storms should be enabled in case a large number of devices wake up simultaneously.	
Enhanced Service Requirements	1.4.1	Connectivity transparency	U-plane interruption time in FDD is 10.5 ms, in TDD is 12.5 ms for both intra-frequency and inter-frequency handovers No specific requirements for inter-RAT interruption time	Connectivity transparency requires: ▪ Seamless connectivity (& intersystem authentication etc.) & mobility interruption time not noticeable by the user (subject to contract) ▪ Inter-system mobility should be controllable by operators ▪ Seamless inter-system authentication, including between 3GPP and non-3GPP RATs	The connectivity transparency refers to the requirement that the user application should be always connected to the RAT or combination of RATs providing the best user experience without any user intervention.
	1.4.2	Security	Security environment to ensure protection of encryption keys and other sensitive information. Mutual authentication, ciphering and integrity algorithms to protect 4G customers.	Security & protection in pervasive & highly heterogeneous environments. Protection despite openness, Big Data, massive connectivity: ▪ Enhanced level of end-to-end security with respect to today's communication systems, in	

				order to protect users' data and detect / prevent / mitigate any possible cyber security attack ▪ Maximize protection against radio jamming of control channels ▪ Provide basic security functions in emergency situations	
	1.4.3	Resilience and High Availability	Not specified	▪ Similar level of robustness as PSTN (99.999%) for the ultra-high reliability, and ultra-low latency use cases. ▪ High resiliency for public safety and emergency communication use cases. - Highly available and resilient for new M2M applications with stringent performance needs. ▪ Remote (self-)healing of equipment should be possible ▪ The system should allow very high service availability	KPI: the chance to have the system properly working.
New Business Model Requirements	1.5.1	Connectivity Providers	Not specified	Allow for evolution of bit pipe model types in both retail and wholesale offerings by exploiting 5G modular architecture and 5G flexibility.	
	1.5.2	Partnerships & XaaS Assest Provider	Not specified	Allow creation of different levels of relationship between operators and application/service providers. Service providers should be able to configure and manage the service via e.g. open API, while operators will have freedom to manage and evolve the network. Ensure operators freedom to manage and evolve the network.	
	1.5.3	Flexible Network	MORAN, MOCN and GWCN	Provide technical instruments to	

		Sharing	mechanisms support network sharing	maximize the overall synergies of network sharing agreements and enable flexible business models that can change dynamically over time.	
Network Deployment, Operational and Management Requirements	1.6.1	Cost efficiency	Not specified	Relates to minimizing the Total Cost of Ownership (TCO) of the network infrastructure for the operators, and the cost of terminals, for any given service offering: ▪ Significant improvement in cost efficiency. ▪ minimize the cost of 5G infrastructure, terminals and operation ▪ offer possibilities for ultra-low cost deployments for very low ARPU areas	The 5G system should enable economically viable deployments to address a very low ARPU (e.g. 100 time lower than current HTC subscribers), by providing a sufficiently low associated TCO.
	1.6.2	Energy efficiency	Not specified	The system has to support the traffic increase of the next decade (possibly in the order of 1000x), with a reduction of the energy consumption of the whole network by a factor 2, leading to a 2000x energy efficiency increase. The energy consumption should be adapted with traffic fluctuation. Energy efficiency gain does not come at the price of degraded performance	KPI: Whole network efficiency: The number of bits that can be transmitted per joule of energy, where the energy is computed over the whole network (including data centres). The system considered for energy efficiency is the whole network (including potentially legacy cellular technologies, RAN and CN and data centres).
	1.6.3	Ease of innovations and upgrade	Not specified	▪ The 5G system has to be flexible enough to enable introducing new access technologies if needed, to be connected to the 5G core network without change to the core network ▪ Similarly, it should be possible	

				to have innovation in the core network and the BSS system with minimal impact on the UE and access network ▪ - The flexibility of the network should allow building a new service much faster than today	
	1.6.4	Ease of deployment	Not specified	▪ Ease to reuse or upgrade from existing network infrastructures. ▪ - Reduced planning, configuration, optimization complexity of whole system.	
	1.6.5	Flexibility & Scalability	Not specified	▪ Flexibility for future-proof evolution; Architecture functional modularity ▪ Ensure HW/SW decoupling within network elements ▪ Scaled service delivery and associated resources utilization (on-demand)	
	1.6.6	Operational Awareness	Not specified	Operator should be aware of users' service and traffic to ensure best user experience.	
	1.6.7	Operation efficiency	▪ CCO: to provide continuous coverage and optimal capacity ▪ MRO: to reduce the no. handover-related RLF ▪ MLB: to cope with unequal traffic load ▪ RACH optimization: to minimize access delays ▪ MDT: to minimize the need of manual drive-tests	Self-acting including efficient plug & play, optimization & recovery Simplified & flexible operation/ management & operational business models. Significant reduction of O&M traffic.	

ANNEX C: REQUIREMENT – TECHNOLOGY MAPPING

			RAN																	Network																				
			Spectrum Access				Radio Link								Radio Access Capacity					Network Flexibility							Efficient/Adaptive Network Resource Usage								Other Enablers					
			Flexible use of licensed spectrum	Integrated unlicensed spectrum	Use of higher frequency bands	Duplex Mode	New waveforms	Advanced multiple access technologies	Radio frame design/ numerology	Massive MIMO and enhanced multi-antenna schemes	Advanced receivers	Interference coordination	Technologies for small packet transmission	UE-centric network	Densification: Small cells/ Ultra-dense networks	Dual connectivity – capacity/ coverage split system design	Enhanced Multi-RAT coordination	Device-to-device communications	Wireless backhauling (e.g. self-backhauling, relay)	Software-defined networking (SDN)	Virtualized mobile core network	Virtualized C-RAN	Flexible split of RAN functions among network	Smart edge node	State-disintegrated core node	Micro-servers	Traffic optimization	Enhanced multi-operator network sharing	Scalable service architecture	Big data	Context-aware/User centered network	Content-optimization and adaptive streaming	Intelligent heterogeneous management	Embedded measurement of network performance	Technologies for massive connectivity	All optical transport network with optical router/switch	Information-centric networking	Mesh Networking	Enhanced fronthauling	Privacy and security
Requirements	User experience requirements	Consistent user experience		X		X		X			X	X		X	X	X	X					X					X			X	X	X	X	X						
		Data rate	X	X	X	X	X		X	X	X	X			X	X	X	X																						
		Latency			X	X	X		X				X	X				X		X			X			X	X									X				
		[High speed] Mobility					X		X	X				X		X														X	X		X							
	Network performance requirements	Coverage capability		X			X		X	X	X	X		X	X	X		X	X									X										x		
		Connection density		X	X			X		X	X	X	X		X	X	X	X								X							X		X					
		Traffic volume density/area capacity	X	X	X	X		X		X	X	X			X	X	X	X									X					X								
		Spectrum Efficiency			X	X	X	X	X	X	X	X			X																									
		Resources and signalling efficiencies					X		X				X	X		X	X	X			X	X	X				X								X		X		x	
	Device Requirements	Operator Control on Devices																																						
		Multi-Band-Multi-Mode Support in Devices	X	X	X	X					X						X																							
		Device power efficiency					X			X		X	X	X	X	X		X															X			X				
	Enhanced Services requirements	Connectivity transparency												X			X													X	X		X				X			
		Context Awareness																												X	X			X						
		Location																																						
		Security																																						X
		Resilience and High Availability														X	X	X	X	X	X	X			X				X				X				X	x		
	New Business Model Requirements	Connection providers																			X	X								X										
		X-as-a-Service																			X	X																		
		Flexible network sharing												X							X	X						X												
		Partnerships and API's																												X										
	Network Deployment, Operational and Management Requirements	Cost efficiency		X			X						X						X	X	X	X	X		X	X	X	X	X			X		X	X	X		x	x	
		Energy efficiency					X			X		X	X		X	X		X			X	X				X							X	X		X				
		Ease of innovations and upgrade																		X	X	X	X		X				X										x	
		Ease of deployment		X	X									X		X	X		X		X	X	X		X													x		
		Fixed-Mobile Convergence																																						
		Flexibility and Scalability	X			X	X		X				X		X	X	X	X	X	X	X	X	X		X	X	X		X			X	X		X		X		x	
		Operation awareness																												X										
		Operation efficiency																		X	X	X	X		X				X	X				X					x	

ANNEX D: TECHNOLOGY BUILDING BLOCKS (PRELIMINARY LIST)

R1 – Spectrum Access

Technology building block name	**Flexible use of licensed spectrum**
Category	RAN
Description	Finer frequency granularity support to exploit any spare spectrum Carrier aggregation to benefit from any spare MHz (Intra-operator / Intra-3GPP) Resource aggregation between higher frequency and lower frequency Efficient utilization of paired and unpaired spectrum Also related to "enhanced multi-RAT coordination"
Specific solutions	Enhanced carrier aggregation schemes
Potential benefits	Improved spectrum utilization
Impact on the network node (s) (e.g. complexity)	More RF to support multiple frequency transmission 1. Devices can be expected to support more than 20 bands out of a possible 100 bands set by 2020 (considering higher frequency bands). 2. Separate antennas and RF chips will be required for frequency bands that are widely separated. 3. Huge implications on the cost (more components, expensive power amplifiers, RF filters, etc.), form factor (many antennas and RF components may have to be fitted) and performance of devices (potential degradation in performance due to insertion losses, antenna structure, coexistence, self-interference, etc.)
Impact on the architecture	Depending on the granularity of control over time, this may require some control interface from some network entity
Maturity	Increasing number of band combinations being supported by carrier aggregation, and work ongoing in 3GPP to support TDD-FDD aggregation.
Missing steps to achieve a good understanding of the technology	Feasibility of a scalable air interface Granularity of frequency control
Challenges	Regulation Avoiding any RF performance degradation

Technology building block name	**Integrated license-exempt spectrum**
Category	RAN
Description	Integrated use of license-exempt spectrum to improve end-user QoE.
Specific solutions	Use of 3GPP RATs in license-exempt spectrum (e.g., LAA - Licensed-Assisted Access), tighter integration of 3GPP RATs with local area technologies that use license-exempt spectrum (e.g., Wi-Fi/3GPP proposal from Intel)
Potential benefits	Allows provision of improved data rates while guaranteeing seamless mobility support.
Impact on the network node (s) (e.g. complexity)	Support for multiple bands required at devices and potentially also at nodes depending on deployment strategies. 1. Devices can be expected to support more than 20 bands out of a possible 100 bands set by 2020 (considering higher frequency bands). 2. Separate antennas and RF chips will be required for frequency bands that are widely separated. 3. Huge implications on the cost (more components, expensive power amplifiers, RF filters, etc.), form factor (many antennas and RF components may have to be fitted) and performance of devices (potential degradation in performance due to insertion losses, antenna structure, coexistence, self-interference, etc.)
Impact on the architecture	Some PHY features (e.g., LBT (listen before talk), DFS (dynamic frequency selection), TPC (transmit power control)) are needed for the RAT to operate on license-exempt spectrum. Some impact on anchor/ aggregation points and policy management entities.
Maturity	Some trial results have been reported for LTE operation on license-exempt spectrum. Some experience with Wi-Fi integration available.
Missing steps to achieve a good understanding of the technology	Coexistence with other technologies operating on license-exempt spectrum needs to be better understood. Impact of transmission conditions (e.g., intermittent transmission due to LBT) on advanced L1/L2 techniques like MIMO.
Challenges	Mechanisms to ensure fair-play w.r.t license-exempt spectrum use between different technologies and operators (i.e., definition and implementation of incentive-compatible spectrum etiquette) need to be developed.

Technology building block name	**Use of higher frequency bands**
Category	RAN
Description	Use of high carrier frequencies (6 GHz and above) for radio access.
Specific solutions	Centimetre waves, millimetre waves
Potential benefits	High frequency communications enable high peak and average data rates as well as low latency in specific scenarios, such as indoor deployments and dense urban areas where high capacity is required. High frequency communications could be seen as a supplement for existing traditional cellular bands. With the combination of higher frequency transmission and massive MIMO, the very narrow beam could facilitate management of intra- and inter-cell interference and enforce the multiplexing. The use of higher frequency bands could provide consistent user experience: 1. Use beaming tracing or fast beam switching 2. High gain beamforming 3. Dual connectivity to an anchor (low) frequency layer
Impact on the network node (s) (e.g. complexity)	The use of high frequency communications involves the introduction of different enabling component solutions at RF, PHY and MAC and antenna levels. Specific solutions are under investigation in particular to split the beamforming tasks between PHY and RF. Different antenna types are under investigation for both UEs and BS. Huge implications on the cost (more components, expensive power amplifiers, RF filters, etc), form factor (many antennas and RF components may have to be fitted) and performance of devices (potential degradation in performance due to insertion losses, antenna structure, coexistence, self-interference, etc)
Impact on the architecture	It may imply specific deployment methodologies. And improvements of interworking capabilities would potentially impact the architecture
Maturity	Propagation characteristic and channel model are not fully understood. Cost-effective transceiver architectural solutions are still under study. There is no fully agreement about the communication scenarios where high frequency communications is likely to make an impact, especially for MMW. For example, current studies are targeting both outdoor and indoor scenarios. MMW has already been standardised in IEEE 802.11ad, although its main use is for niche applications, like remote desktop. The use of 802.11ad for broader applications has to be verified. Availability of measurements results in higher frequency bands: 802.11ad has some measurement results at 60GHz. NYU and University of Austin also has some measurement results. EU Projects Miweba and MiWave are studying these aspects.
Missing steps to achieve a good	Further study and selection of high frequency bands Explore the typical scenarios and deployment

understanding of the technology	Understand the effect of blockage and of real-world propagation effects Design enabling P2P and P2MP multi-antenna solutions, with reasonable cost and energy consumption figures System design, in particular to understand synergies between low-frequency and high-frequency bands (e.g. air interface design, physical frame, scheduling, channel estimation…) 1000 x capacity analysis Cost analysis
Challenges	Propagation, air interface design, cost-effective transceiver design, integration with low-frequencies, cost-effective dense deployment. Service experience may not be consistent with highly fluctuating quality levels due to propagation, movement of terminal and difficulty to cover large area Indoor coverage from outdoor base station: Depends on the type of building and the building materials. Generally, with high gain beamforming, low range GHz bands can still do outdoor-indoor well. For mmW, the huge penetration losses will make it challenging. Indoor base station deployments may be needed.

Technology building block name	**Duplex Mode**
Category	RAN
Description	1) Full duplex. Simultaneous transmit and receive on the same spectrum and time resources. 2) Flexible duplex mode, e.g., unified TDD/FDD frame structure design. 3) FDD symmetric/ asymmetric DL/ UL bandwidth allocation. 4) TDD with more flexible UL/DL configurations
Specific solutions	
Potential benefits	Significant increase of capacity (up to 2x) with full duplex. Potential lower latency with full duplex. Flexible downlink and uplink scheduling with the DL and UL channel information and interference conditions Integration of backhaul and access for heterogeneous networks in NLOS scenarios Better behaviour of advanced receivers' in TDD modes thanks to stable covariance matrices Channel reciprocity in TDD modes for enhanced multi-antenna solutionsFull duplex could adapt to dynamic DL-UL traffic load with flexible resource allocation Flexible duplex could also minimize the difference between TDD and FDD design, and reduce the complexity of co-platform implements
Impact on the network node (s) (e.g. complexity)	Expected complexity increase due to severe interference situations BS may take different duplex modes, BS coordination or C-RAN type central processing is needed. Different Tx and Rx antennas may corrupt channel reciprocity in TDD systems Tight network synchronization required in TDD as well as inter-operator coordination for coexistence More UL-DL and DL-UL interference could be introduced by flexible duplex mode. Overall, flexible duplex modes will increase cost and complexity (design, overprovisioning, testing, etc)
Impact on the architecture	BS may take different duplex modes, BS coordination or C-RAN type central processing is needed. Distinction between TDD and FDD can be blurred, or completely removed, facilitating a unified but flexible duplex mechanism Full duplex could enable unified structure of TDD and FDD
Maturity	Research topic Companies have already claimed that their self-interference cancellation scheme can be applied successfully to full duplex transmissions There are multiple demos showing self-interference mitigation works well. Little works on Intra cell interference mitigation, and inter cell interference mitigation. Full duplex could evolve based on TDD-FDD carrier aggregation

	schemes, dual/multiple connectivity, eIMTA (Enhanced Interference Mitigation & Traffic Adaptation) etc…. Feasibility of flexible TDD: Feasible, already possible with LTE. Issue is with interference management. Feasibility of flexible FDD: Feasible if dynamicity is not required. Challenging if dynamicity is required. Challenging to design tuneable duplexers requires more advanced interference suppression techniques if duplex distance is variable/too low. Feasible applications of full-duplex: 1. self backhauling 2. mmW backhaul 3. small cell backhaul 4. relay 5. low-latency C-Plane backhaul for CoMP Current Full duplex technology not mature enough for extreme self-interference cancellation required for general purpose FD. Some tailored solutions relying on combinations of more antenna isolation and advanced interference suppression may be feasible to enable the mentioned applications Self-interference cancellation level in FD: 80dB - 100dB (analog + digital) SoTA. Target is ~120dB for a pico cell. ~140dB for macro cell
Missing steps to achieve a good understanding of the technology	To assess performance and complexity of interference cancellation schemes Interference mitigation in Network with full duplex needs more study: intra/inter cell, inter user interference (UL to DL, UL to UL, DL to DL, DL to UL) Joint UL and DL power control, scheduling, reference signals design Frame structures To assess capacity gain in real field To assess component cost/performance benefit
Challenges	Interference management and cancellation Component cost Foreseen issues for full duplex: 1. Tx/Rx antenna isolation and signal rejection 2. Self-interference cancellation

R2 – Radio Link

Technology building block name	**New Waveforms**
Category	RAN
Description	Alternative to or enhancements of OFDM. Need to be explored and seen if they would allow step-improvement or not, or used for specific requirements/ scenarios
Specific solutions	Generalised Frequency Division Multiplexing (GFDM), Filter-Bank Multi-Carrier (FBMC), Faster Than Nyquist (FTN), Wave Amplitude Modulation (WAM), Sparse Code Multiple Access (SCMA), Filtered-OFDM (F-OFDM), Universal Filter Multi-Carrier (UFMC). Enhancements of OFDM (modified Cyclic Prefix) and single carrier (SC) FDM.
Potential benefits	Enable steeper spectrum roll-off and bands with difficult constraints Reduced peak to average power ratio (energy efficiency, cheaper devices, especially M2M) Enable framework with low latency Address M2M specific requirement, especially for low data rates Higher spectral efficiency (but still not properly quantified) Enable synchronous-constrained applications (e.g., COMP, D2D, MMC) High transmission range in higher frequency bands Combination with Massive MIMO in higher frequency bands
Impact on the network node (s) (e.g. complexity)	Terminal and BTS RF and baseband components MIMO + COMP design
Impact on the architecture	Minor. However, some specific waveforms may address specific needs and may only be applied there.
Maturity	Investigated in research projects/start-ups, comparison with specific solutions ongoing
Missing steps to achieve a good understanding of the technology	Comparison with LTE Rel. 12 and M2M systems including system-level impact. Quantification of latency including synchronisation and detection. Is there consensus for new waveform(s)? 1. No. OFDM, with some enhancements (e.g., numerology) sufficient for up to 20-30 GHz frequency range and most use cases. More thoughts needed for above 30GHz 2. Yes. MTC use case calls for new waveform (e.g., narrowband discrete single carrier, FBMC). 3. Yes. SCMA provides new properties to flexibly support diversity of use case 4. Yes. UF-OFDM to replace OFDM. Improve spectral properties and improved robustness to time and frequency. Different waveform may be needed for mmW
Challenges	OFDM is well-established, , the following challenges need to be addressed: ▪ Integration with MIMO, Multiple Access, HARQ etc

	▪ Robustness to hardware impairments such as carrier frequency offsets to deal with low-cost devices, which may feature low-cost oscillators and radio-frequency front-ends. ▪ Support for frequency selective link and rank adaptation.

Technology building block name	**Advanced multiple access technologies**
Category	RAN
Description	Advanced multiple access technologies should provide higher network spectral efficiency; the performance gap between cell-centre and cell-edge users could be reduced, and the number of simultaneous (access) users could be increased. Non orthogonal multiple access (NOMA) scheme efficiently exploits the channel gain difference among/between users to achieve high spectral efficiency. In NOMA, multiple users can transmit signals at the same spatial-time-frequency resource during uplink transmission, or the signals of multiple users can be transmitted by eNB at the same spatial-time-frequency resource during downlink transmission. To obtain multi-user multiplexing gain, advanced interference cancellation should/must be carried out/implemented on receiver side. Additionally, power allocation and multi-user scheduling are needed at the transmitter side.
Specific solutions	Non-linear interference cancellation method Power allocation User grouping
Potential benefits and requirements addressed	▪ Provides high network spectral efficiency ▪ Access more users / Number of simultaneous users increased ▪ Improves the cell-average throughput and cell-edge user data rate.
Impact on the network node (s) (e.g. complexity)	Introduce an interference cancellation baseband unit at the receiver side, and increase the receiver's complexity
Impact on the architecture	Minor
Maturity	Study on SIC receiver is done in Rel-12 for inter-cell interference cancellation Started in some 5G project such as METIS, 5GNOW
Missing steps to achieve a good understanding of the technology	Need to demonstrate its performance gain for future practical application
Challenges	Introduce a PHY and MAC challenges to the existing network topology

Technology building block name	**Radio frame design / numerology**
Category	RAN
Description	Design of the radio frame should be carried out in such a way that all the requirements are met. We note that different requirements may lead to different radio frame designs; so it has to be verified whether a single radio frame can meet all the requirements. Optimized/Flexible design of the frame structure and modulation parameters.
Specific solutions	N/A
Potential benefits	Lower latency, energy saving, overhead reduction, increased robustness, elasticity, reconfigurability.
Impact on the network node (s) (e.g. complexity)	Complexity, transceiver design (e.g., ADCs/DACs, RF), synchronisation, channel-state estimation quality, signalling overhead, transmission range, channel aging, etc…
Impact on the architecture	No major impact on architecture.
Maturity	Specific solutions on the radio frame depend on other building blocks, still under investigation.
Missing steps to achieve a good understanding of the technology	Performance evaluation of different radio frame designs for different frequency regimes and different deployment/use scenarios
Challenges	Achieving a single/scalable design suitable for different frequencies and use cases

Technology building block name	**Massive MIMO and enhanced multi-antenna schemes**
Category	RAN
Description	Use of a large number of antenna elements within one antenna panel or across multiple panels or locations. Each (or a group of) antenna element has an independent RF-chain, which can enable advanced beamforming (2D, 3D or full dimensional beamforming), MU-MIMO, multiplexing, diversity and interference mitigation. Spatial Modulation approach: a single radio frequency used in a multi-antenna context
Specific solutions	Massive MIMO as extension of LTE-advanced or a new design: 1. LTE framework for MIMO could be used/extended for TDD and also for accommodating ~16 antennas/ <6GHz. (under 3GPP discussion) Improvements needed in CSI feedback and sounding schemes. No radical change from LTE needed. 2. New/radical designs needed for both FDD and TDD with massive MIMO when antenna number is tremendously extended (e.g., 64 x 64). Improvement areas: configurable reference schemes, direct compression channel feedback, reference signals, link adaptation, control channels. Also implications of massive MIMO on architecture need to be addressed (handover, signalling, info exchange)
Potential benefits	Large number of simultaneously served users Large array gains High directivity High spectral efficiency (SE) Interference mitigation capabilities Optimality of low-complexity signal processing High-rise buildings' coverage from outdoor eNB, which reduces the difficulty of indoor system deployments Low power and low cost RF components High energy efficiency (EE) Larger region of simultaneous EE and SE increase in the EE-SE curve Artistic placement of antenna elements in DIY manner for potential business models
Impact on the network node (s) (e.g. complexity)	Major impact at the RF level in the BTS. In general, except when low load for which elements can be switched off, all elements composing the antenna panel must be active to allow for the highest degree of beam steering and interference mitigation flexibility. Size of array strongly depends on frequency band used Large amount of RF chain will cause the complexity of RRU and baseband processing Large amount of data exchanges through interface such as CPRI Lower power requirement of each port will lower the cost of PA and size of filters
Impact on the architecture	Impact on deployment. Needs to be seen in conjunction with COMP. Needs to be seen how to deploy in Heterogeneous Networks. Needs to be seen how to cover high rise buildings. May possibly alleviate the burden of adding small cells in the coverage

	of macro BS with massive MIMO, via beamforming to the region. The inband wireless backhauling can facilitate easier deployment. Need of inter-cell coordination scheme for massive MIMO: 1. For ensuring continuity of coverage, some degree of coordination may be needed. E.g., with C-RAN 2. Massive MIMO is interference resistant (more so as the number of antennas increase) due to selective beamforming and power control. In practical systems with limited number of antennas, flashglight interference effects could occur, but this will not have a significant impact as the system tends to be noise-limited, rather than interference limited at higher frequencies. 3. Codebook restriction, coordinated beamforming and coordinated scheduling all on different timescales could help to address any potential interference issues. 4. Gain dependent on deployment and scenario
Maturity	A lot of theoretical studies have been carried out to evaluate the advantages and challenges of this technology. Current studies are focused on considering the impact of real-world effects. Beam forming in backhaul application more mature 3GPP provides 3D channel modelling of 3D UMa (Urban Macro) and 3D UMi (Urban Micro). But other use cases channel modelling is not ready, such as high-rise buildings and indoor (e.g. office, stadium, malls…) Practical antennas number and array configuration in massive MIMO by 2020 + related performance: 1. Difficult to answer. Depends on deployment scenarios 2. Arrays with ~100 elements feasible at BS side. ~10 at UE. Performance depends on scenario. Need careful study. 3. 256 element arrays (with 64 RF chains) at BS at 2Ghz. Coverage depends on control channel design 4. 64-128 elements at 2GHz. 256 elements at 6GHz 5. 128 elements arranged in 2 8x8 arrays
Missing steps to achieve a good understanding of the technology	▪ Channel estimation (e.g. pilot contamination and pilot design, channel aging) and channel state information feedback enhancements ▪ Application to FDD systems ▪ RF cost and energy consumption ▪ New antenna deployments, new cellular structure ▪ Coverage enhancement for (common) control channels ▪ Utilization of TDD channel reciprocity ▪ Environment friendly antenna configuration and deployment ▪ Efficient calibration method for a large number of RF chains ▪ Hybrid beamforming structure, where the number of transceivers can be much smaller than the number of antennas ▪ Control channel and data channel coverage ▪ Distributed beamforming algorithms to avoid centralised

	computation of coefficients. ▪ User scheduling. ▪ Cooperation: how can massive arrays cooperate? ▪ Distributed array, or centralised array? ▪ Beamforming vs. spatial multiplexing dilemma: the former mainly applicable in LOS, the latter requiring channels with many degrees of freedom (not always encountered in practice) ▪ Coexistence with CRAN very challenging because of the extreme data rates in the fronthaul
Challenges	See above.

Technology building block name	**Advanced receivers**
Category	RAN
Description	Ability to cater/cope with intra-cell and inter-cell interference at the receiver side. Extra signalling from the network in order to help advanced receivers when coping with interference
Specific solutions	Evolutions from SIC/PIC; intra-subcarrier equalization; inter-symbol interference (ISI) cancellation for Faster-than-Nyquist; intra-cell interference cancellation for non-orthogonal waveforms and non-orthogonal multiple access (NOMA); Inter-Carrier Interference (ICI) and ISI cancellation for CP-free OFDM
Potential benefits	Increased simplicity at the network side (relaxed coordination and synchronisation) by allowing a controlled amount of interference at the RX side. Extra signalling to be considered at the network side in order to aid advanced receivers in cancelling inter-cell interference
Impact on the network node (s) (e.g. complexity)	Less complex transmission methods (less need of interference coordination/cancellation at the transmitter side) Additional signalling mechanisms to help receivers from the network side
Impact on the architecture	More simplified architecture if cell cooperation is less required Additional/different reference symbols and feedback may be needed in order to support better terminal receivers
Maturity	Very immature when considering non-orthogonal waveforms Proprietary advanced receivers are very mature, but terminal classes/requirements don't honour better receivers
Missing steps to achieve a good understanding of the technology	Need to agree on the potential of non-orthogonal waveforms and/or multiple access schemes prior to dealing with extra interference
Challenges	Extra processing power at the RX side; possibly reduced performance About spectral efficiency/latency trade off: 1. Processing latency at the receiver is seen as insignificant/negligible compared to other sources of latency for most use cases. May become significant for mission critical use cases such as tactile internet. 2. Hardware design with higher processing capabilities can improve latency but will come at a higher cost. 3. Linear type of advanced receivers only adds insignificant delay. For non-linear types (e.g., SIC receivers), processing delay depends on implementation (soft or hard decision, hardware, etc.)

Technology building block name	**Interference coordination**
Category	RAN
Description	To avoid interference via information exchange between schedulers on network side.
Specific solutions	Inter-Cell Interference Coordination – Rel.8 Coordinated Multi-Points (CoMP) – Rel. 11 & 12 In addition: definition of coordination messages on standardized inter-eNB interfaces for multi-vendor deployments, and/or definition of a central coordinating node, and/or introduction of the beams elevation dimension into the coordination parameters.
Potential benefits	Enhanced reception performance and reduction of the number of dropped calls A better network resource utilisation by using least loaded base station in CoMP process. An increased received power by joint reception from multiple base stations A reduced interference level by utilising constructively the interference rather than destructively.
Impact on the network node (s) (e.g. complexity)	In one hand CoMP needs a very low level of latency in other hand multiple site reception and transmission add significantly to any delays (communication between different sites). The maximum performance is obtained with centralized scheduling of a quite large number of base stations. That increases the backhaul traffic.
Impact on the architecture	As different sites may be connected together C-RAN is proposed/may be applied.
Maturity	CoMP has been the focus of many studies by 3GPP for LTE-Advanced as well as the IEEE for their WiMAX, 802.16 standards. For 3GPP some CoMP schemes were introduced in R11.
Missing steps to achieve a good understanding of the technology	Defining properly the CoMP candidates
Challenges	Backhaul capacity

Technology building block name	**Technologies for small packet transmissions**
Category	RAN, CN
Description	There is a wide variety of small packets transmission schemes with different QoE (Quality of Experience) for both M2M and H2H, e.g. 1) Periodical keep-alive packets 2) Bursty Instant Messages 3) Real-time critical message delivery … The above small packet transmission may cause frequent RRC (Radio Resource Control) transitions and result in network signalling congestion. And moreover, the current RRC transition may introduce extra delay and can't satisfy the real-time requirement of some small packets transmission. 3GPP is looking for signalling optimization of small packets transmission. Enhanced mechanisms need to be devised for 5G.
Specific solutions	Radio/CN signalling reduction needs to be investigated further, e.g. UL Data Transmission integrated with Random Access. Some new scheduling without too much signalling overhead should also be studied, e.g. Periodic scheduling of Keep-Alive. Contention based access considering network loading and trade-off collision probability is also one of the promising candidate technologies. Furthermore, a mechanism for rapid RRC state transition needs to be devised.
Potential benefits	Significant signalling reduction and more rapid message delivery
Impact on the network node (s) (e.g. complexity)	Not a big impact on network node design. New signalling procedures, new scheduling mechanism and rapid RRC transition is required.
Impact on the architecture	No impact
Maturity	3GPP has carried out some initial study on this and more research is needed.
Missing steps to achieve a good understanding of the technology	Need to agree on what kind of small packets with specific QoE needs to be optimized individually.
Challenges	To be explored

Technology building block name	**UE-centric network**
Category	RAN and CN
Description	Cellular designs have historically relied on the role of “cells” as fundamental units within the radio access network. However during the last years different trends emerged that call for a disruption of this concept. These trends include [BHLMP 14]: - Cell densification, with base stations using very different transmit powers and antenna configurations. We note that while cell densification is supported in LTE, the current architecture hasn’t been designed to *natively* support small cells. - the need of additional spectrum will lead to the coexistence of frequency band with very different propagation characteristics, from very low to very high frequencies - RAN virtualization will lead to a decoupling between a node and the hardware allocated to handle the processing associated with this node. Hardware resources in a pool, for instance, could be dynamically allocated to different nodes depending on metrics defined by the network operators. - New service classes, requiring the content and/or the application server and/or some core network functions to be placed nearer to the users. This calls for an elastic architectural design, where the topology of the network is adapted to the specific services deployed. - The use of smarter devices could impact the radio access network. In particular, both D2D and smart caching call for an architectural redefinition where the centre of gravity moves from the network core to the periphery (devices, local wireless proxies, relays) Based on these trends, the UE-centric network vision [BHLMP 14] calls from an evolution of the “old” cell-centric architecture into a device-centric one: a given device (human or machine) should be able to communicate by exchanging multiple information flows through several possible sets of heterogeneous nodes. In other words, the set of network nodes providing connectivity to a given device and the functions of these nodes in a particular communication session should be tailored to that specific device, service and session. Mobile core network can be flexibly sliced into several overlay core network serving for different type of users, UEs or use case, like M2M users and so on. [BHLMP 14] F. Boccardi, R. W. Heath, A. Lozano, T. Marzetta and P. Popovski, *Five Disruptive Technology Directions for 5G,* IEEE. Comm. Mag. Feb. 2014.
Specific solutions	The following technical solutions are enabling components for the UE-centric network: native support of uplink-downlink splitting, native support of control and data separation, elastic association of cells to users, elastic association of cells to hardware resources, native support of CoMP, use of distributed content e.g. at the edge and at the UE, function block design in finer granularity and others.

Potential benefits	Enabler for new services, reduction of costs, boost in performance for traditional service. 1. Inherent support for multi-point Tx and Rx schemes, carrier aggregation, UL/DL splitting 2. Robust mobility support 3. Improved UE energy consumption, lower latencies, lower overhead, flexible network deployment
Impact on the network node (s) (e.g. complexity)	See discussion above 3. No major impact on UE cost
Impact on the architecture	See discussion above. 1. Changes to traditional handling of mobility (handover, RA, paging) 2. Changes to broadcast channels, inter-node interfaces, RRM
Maturity	Some of the component technologies are quite mature, some others require further studies. Interesting area with good potential which deserves further investigation for 5G
Missing steps to achieve a good understanding of the technology	
Challenges	

R3 – Radio Access Capacity

Technology building block name	**Densification: Small Cells / Ultra-dense networks**
Category	RAN
Description	Densification of the network via deployment of small cells
Specific solutions	
Potential benefits	Capacity increase Also coverage enhancement (e.g. indoor) & greater uniformity
Impact on the network node (s) (e.g. complexity)	Great impact especially with today's perspective. The roadmap (particularly Rel-11/12 onward) to continuously address signalling, phantom and lean carriers, interference management, plug and play configuration, centralized/distributed macro/micro coordination, non-ideal backhaul, etc.
Impact on the architecture	Backhaul to be provided for small cells (Need to be greatly simplified and flexible in configuration & optimization/operation, & efficient in resource allocation and energy efficiency) Need embedded sustainability too with great number of nodes. May need a clean slate approach to support e.g. phantom cell concept, UL/DL separation…
Maturity	Ongoing: Defined with initial deployments, continuous maturity (3GPP, SCF, NGMN). Emerging phase considers SON, backhaul, multiband/CA considerations. However, the need for densification e.g. for 2020 vision requires more to enable capacity in an efficient and flexible (& potentially ad-hoc and pervasive add/delete & on/off) manner.
Missing steps to achieve a good understanding of the technology	The link between emerging (Rel-12, NGMN project, SCF releases) and the maturity vision indicated above, from a picture of add-on cells to low-power macro-cell capacity as needed →close to user. Enabling the densification picture that looks very complicated (& crowded) today.
Challenges	Backhaul Interference management Please also see notes under impacts, maturity and missing steps.

Technology building block name	**Dual Connectivity - Capacity/coverage split system design**
Category	RAN
Description	A user terminal can be associated to two different base stations for uplink and downlink, respectively.
Specific solutions	Not yet available.
Potential benefits	This is suitable in HetNets, where there are noticeable asymmetries between cells of different sizes. This type of connectivity has the aim at optimising the quality of the uplink connection when a device is in range of both a small cell and a macro cell. Advantages with respect to range extension include optimisation of the uplink without performance drop in the downlink. In addition, terminal transmission power can be backed off if associated to a small cell placed in physical proximity. Further benefits for asymmetric services (e.g., M2M…) Improved spectral efficiency in UL and DL, Load balancing, Interference coordination, reduced UE power consumption
Impact on the network node (s) (e.g. complexity)	Major impact on node design, e.g., different nodes for uplink and downlink.
Impact on the architecture	Major impact on architecture.
Maturity	Research on this topic just started. Views 1. Low latency HARQ, power control and CSI feedback favour co-located DL/UL points. Higher level control and mobility management can be split. Added cost to provide UL channel once DL channel is provided is very incremental 2. For same sized cells, benefits of UL/DL splitting may not justify additional standardization complexity, given that UL and DL scheduling is already done independently in SoTA 3. UL/DL splitting may provide some gains for HetNets when cells are of different sizes
Missing steps to achieve a good understanding of the technology	Further performance evaluation needed, further assessment of impact on system design is needed (e.g., frame design, channel state estimation, control channel, etc…)
Challenges	A user terminal can be associated to two different base stations for uplink and downlink, respectively. CSI acquisition, HARQ, user identification, UL power control, coordination of traffic flows across different paths, UE complexity, ideal backhaul

Technology building block name	**Enhanced multi-RAT coordination**
Category	RAN
Description	Facilitate uniform and converged management of multiple RATs (LTE, 5G, Wi-Fi) by re-designing radio/ network functions and decoupling the control and user planes
Specific solutions	Multi-RAT coordination and convergence, Multi-connection and Multi-transmission, RAN functionality evolution including data cache, smart service and contents distribution and aggregation, etc. Control and user plane decoupling
Potential benefits	Improved network resource/ operation efficiency Improved user experience
Impact on the network node (s) (e.g. complexity)	A new network logical entity may be introduced, which enables efficient coordination to facilitate multi-RAT access, in an agnostic manner and seamless to the end user 1. Reduction of legacy interworking requirements on the network side 2. More components (network functions) to manage
Impact on the architecture	A new network logical entity may be introduced
Maturity	A related study item in 3GPP Rel-13
Missing steps to achieve a good understanding of the technology	The benefits and drawbacks of introducing such a new network entity to manage multiple RATs and decoupling of the control and user planes need to be fully understood. Impact of such coordination entity on network assistance (e.g., signalling, scheduling coordination, considering device transmission power limitations, etc.) needs to be fully understood.
Challenges	Introduction of a new network logical entity will bring challenges to the existing network topology

Technology building block name	**Device-to-Device communications**
Category	RAN
Description	Two devices can directly exchange data without having to route it through a network. This is already possible with state-of-the art technology, but communication can only happen after manual pairing, and this is often platform-dependant.
Specific solutions	D2D links can be established on cellular spectrum (i.e., in-band), or license-exempt spectrum (i.e., out-band), and the network can be involved (controlled D2D) or not (autonomous D2D) in the establishment and maintenance of the connection. The solutions that are being discussed now are: ▪ D2D for Public Safety in 3GPP Rel. 12. ▪ LTE-Assisted Wi-Fi Direct. ▪ Different versions of LTE-Direct.
Potential benefits	This can be an opportunity for future cellular systems to enhance their performance in terms of: Network off-loading, spectral efficiency, throughput, fairness, coverage extension, latency and power-saving. In addition, it gives the possibility of carrying out emergency calls in out of coverage areas.
Impact on the network node (s) (e.g. complexity)	For in-band controlled D2D: devices have to be equipped with the new capability (currently not available). For out-band controlled D2D: devices are already WiFi-capable. The impact on the network depends on the approach: if discovery is carried out over the top, some components need to be added to the network infrastructure (i.e., D2D server), and possibly new signalling schemes need to be introduced to allow the network to perform peer discovery, connection establishment, service continuity etc…
Impact on the architecture	Security architecture, service delivery, legal intercept, privacy, billing etc.
Maturity	There is a need to quantify the benefits, and find the right use cases.
Missing steps to achieve a good understanding of the technology	Analysis needs to be done to see what the exact benefits to a network/cloud approach are.
Challenges	Legal intercept obligations, policy and charging enforcement, added value for the customers as compared to e.g. Wi-Fi direct

Technology building block name	**Wireless backhauling (e.g. self-backhauling and relay)**
Category	RAN/CN
Description	Replace some of the existing cabled backhaul links by 5G wireless links Backhaul needs to be integral part of the architecture. What interfaces need to standardized or can be left proprietary needs to be defined
Specific solutions	For static ultra-dense networks of nodes: mm-waves For moving nodes (vehicles): < 6 GHz, and massive MIMO beamforming with predictor antennas on vehicles. For static nodes in rural environments: < 6 GHz and massive MIMO beamforming
Potential benefits	Avoid the coming vehicular data tsunami overloading of the operators' spectrum and networks.
Impact on the network node (s) (e.g. complexity)	The support of moving nodes requires additional complexity and PHY/MAC to provide robust performance.
Impact on the architecture	
Maturity	Research activities have started.
Missing steps to achieve a good understanding of the technology	Proof of concepts of mm-waves, Massive MIMO and predictor antennas.
Challenges	To design low complexity solutions.

N1 – Network Flexibility

Technology building block name	**Software-Defined Networking**
Category	Network
Description	Programmable network with centralized logically abstracted control, separated from a flow-based data/forwarding plane, like P-GW/S-GW and so on.
Specific solutions	OpenFlow-based SDN is the first “standard” out of the SDN body, Open Networking Foundation. Wireless and mobile working group of ONF is working on SDN-based mobile packet core network. Need to consider input from ETSI working groups.
Potential benefits	Efficiency, flexibility, programmability and therefore broad ecosystem with fast pace of innovations and upgrade, granular view/management, and energy efficiency. Control plane and data/forwarding plane of existing functions (like S-GW and P-GW) is further separated. That contributes to the centralization of control plane of network functions, the distribution of data plane functions to localize data traffic and more flexible steering of data traffic according to operator policies.
Impact on the network node (s) (e.g. complexity)	Network nodes need to be configurable. Overall node complexity is reduced by separated control plane.
Impact on the architecture	Simplification, flexibility and higher efficiency. Separation of logically abstracted control plane and forwarding data planes. Open APIs between app (& business processes) plane and control planes (northbound), control & data planes (southbound), and between (sub-controller) domains (East-West). Application definitions realizing current architecture functions like mobility management of core network are open.
Maturity	Concept since 2008 with some deployments especially in IT domain. A great deal of R&D, activities and initiatives in this area. Specification by ONF, first of their standards published based on OpenFlow. (southbound API and flow logic based on flow tables, matching entries and associated actions) The extension to telecom seems to be at early stages and has much room to mature. Wireline/fixed environments are closely explored (e.g. to collapse CO’s).
Missing steps to achieve a good understanding of the technology	SDN is a package initially defined in the IT domain. More needs to be done to extend it to the telecom/mobile environment - for example, identification of key issues in implementing a SDN based EPC framework to support functional separation between control and user planes for SGW and PGW.
Challenges	Deeper understanding required and clearer/more detailed input to ecosystem/standards

Technology building block name	**Virtualized Mobile Core Network**
Category	Network
Description	Software based functionality abstracted from common pool of hardware. Enables mobile core network elements as virtualized functions decoupled from specialized hardware, managing function and resources more flexibly and intelligently Virtualization platform can provide open APIs to management functions utilizing shared resources.
Specific solutions	NFV
Potential benefits	▪ Efficiency in cost, end to end energy usage efficiency by maximizing pool of common physical resources. ▪ Flexibility in deployment, resource allocation and management ▪ Broader ecosystem with faster realization of innovations and upgrades ▪ An enabler for end to end virtualization as part of a service (or instance defined), and for requirements such as use-case & context-specific resource allocation
Impact on the network node (s) (e.g. complexity)	Changes the concept of network nodes. Group of virtualized functions can be run in one data centre. One virtualized function can be instantiated at different data centres with different parameters. Positive impacts expected such as flexible pooling arrangements, sharing out of a cloud of resources as well as more granular yet simplified management
Impact on the architecture	Great impact on architecture given that functions become virtualized entities decoupled from hardware. Hardware can be shared and possibly non-specialized (e.g. standard processors, switches, storage). The resource environment can be sliced for independent usage for testing and proof of concept next to live networks. A s/w based forward graph orchestrating different functions from radio, core and services can define a communication path.
Maturity	ETSI ISG work on NFV. 3GPP just asked for contributions on what it needs to engage in related to Virtualization. Vendors are working on virtualization in general, including virtualized EPC and IMS implementations with several commercial solutions available today, notably using COTS IT hardware. Operators are exploring and generally in consensus on virtualization as a future direction, with some already reporting on their plans. Significant work is still needed in this area to move towards maturity.
Missing steps to achieve a good understanding of the technology	Although much has been discussed and explored in telecom/mobile domains, the technology needs to be further defined, tested and proven - for instance, how to adopt the service chain concept for C/U-plane telecom functions.
Challenges	Deeper understanding, modelling and more clear/detailed input to ecosystem/standards.

Technology building block name	**Virtualized C-RAN**
Category	RAN
Description	To enable C-RAN as virtualized functions decoupled from hardware - a step beyond pooling and C-RAN
Specific solutions	Virtualization technology borrowed from IT industry
Potential benefits	▪ Efficiency in cost and end to end energy usage ▪ Flexibility in deployment, resource allocation, and management ▪ Broader ecosystem, faster innovations and upgrades ▪ Works well with end to end virtualization as part of a service or instance defined, and with important requirements such as use-case and context-specific resource allocation ▪ Potential enabler for other 5G technologies, especially interference-cancellation/cooperation technologies such as CoMP ▪ Potential to better support multi-RAT and multi-RAN coordination
Impact on the network node (s) (e.g. complexity)	Positive impact is expected in terms of pooling, resource sharing and flexible use as well as more granular yet simplified management. Further exploration is required.
Impact on the architecture	Great impact on architecture given that functions become virtualized entities decoupled from hardware. Hardware can be shared and possibly non-specialized (e.g. standard processors, switches, storage). The resource environment can be sliced for independent usage for testing and proof of concept next to live networks. A s/w based forward graph orchestrating different functions from radio, core and services can define a communication path.
Maturity	ETSI ISG work on NFV as well as NGMN RAN-EV project exploration is in progress. 3GPP recently asked for contributions on what it needs to engage in on the topic of virtualization. Vendors are working on virtualization in general. Already some operators reported successful large-scale C-RAN deployment or field trials. Centralization is not the major obstacle to overcome. WDM-based fronthaul solution is mature enough. Exploitation of virtualization technology with respect to RAN implementation has just started.
Missing steps to achieve a good understanding of the technology	Virtualization in general, to expand from IT concepts into telecom/mobile and to radio in particular (as service and core architectures are considered closer to web and closer to IT experience).
Challenges	Virtualization implementation to meet critical real-time processing requirements, including load balancing and fronthaul latency requirements. Deeper understanding, modelling and more clear/detailed input to ecosystem/standards.

Technology building block name	**Flexible Split of RAN Functions Among Network Nodes**
Category	RAN / L1 and L2
Description	Ability to perform centralization/decentralization of L1/L2 RAN protocols (or subparts of them) according to specific needs and flexibility in assigning protocol functions to RAN nodes.
Specific solutions	Separation of control and user planes; independence of the RAN protocols and the network nodes that run them.
Potential benefits	▪ Centralization would save processing node resources whenever possible, and allow for L1/L2 cooperation among multiple nodes. ▪ De-centralization would save backhaul resources by limiting the area of the network where L1/L2 signals would flow. ▪ Centralization/de-centralization should be adaptive according to the particular scenario.
Impact on the network node (s) (e.g. complexity)	Large impact on network node design, requiring the decoupling of HW/resources and protocols. Would be transparent for the users (devices)
Impact on the architecture	Architecture should be flexible enough to cope with the two extreme cases above
Maturity	Still very immature for L1/L2 functions; more mature for L3 and applications
Missing steps to achieve a good understanding of the technology	Making the network flexible enough to cope with both centralized and distributed architectures. Maturity of GPP solutions for L1.
Challenges	Required processing power of generic hardware for performing intensive L1 tasks. High bandwidth requirements for backhaul network in centralized scenarios at PHY level.

Technology building block name	**Smart Edge Node**
Category	Architecture
Description	A node at the edge of the network (e.g., base station, small cell or even terminal) can actively carry out some of the core network functionalities or additional services (example: context-aware dynamic caching)
Specific solutions	
Potential benefits	Pushing network intelligence to the edge has the potential to improve system latency, customer experience and resilience.
Impact on the network node (s) (e.g. complexity)	Edge nodes will become more complex as they have to emulate parts of the core network, or serve as reliable caching units. A virtualized software-based network can offer scalability to move functionality to the core or edge based on requirements. For many services and network function, the smart edge node can consist in a mini-datacentre using general purpose servers. New applications can then be easily deployed as virtual machines, with hardware resources shared with other applications.
Impact on the architecture	This will blur the boundary of CN and RAN. Part of the core network functions will be distributed towards the edge, part will still be centralized.
Maturity	Advanced developments in the ecosystem, including proof-of-concept trials. Some small cell products already support the flexible installation of operator deployed functions.
Missing steps to achieve a good understanding of the technology	
Challenges	Creating a flexible, reconfigurable and scalable solution. Keeping deployment cost low. Backwards compatibility with legacy systems. Coordination of multiple Radio Access Technologies.

Technology building block name	**State-disintegrated Core Node**
Category	Network
Description	State of a core node is separated and kept in a remote database
Specific solutions	
Potential benefits	▪ Reliability, faster recovery, low-cost redundant system design in virtualised networks. ▪ Reduced redundant state information in multiple core nodes.
Impact on the network node (s) (e.g. complexity)	Major. Have the potential of reducing complexity of core nodes. API might be needed (depending on the actual implementation) to fetch/update the state information in the remote database.
Impact on the architecture	Major. Session/user state in core nodes needs to be stored in a remote location leaving core nodes to perform processing duties only. If one node fails, another node can simply look up the states in the remote database and instantaneously take over from the failed node. Reliability/redundancy of the database system storing the states is crucial. Latency impacts associated with remote state retrieval must be minimized.
Maturity	Industrial labs have reported working prototypes.
Missing steps to achieve a good understanding of the technology	Reliability of the backend database (DB) system storing states. State exchange load in between core nodes and the DB. Cost benefit
Challenges	See above

Technology building block name	**Micro-servers**
Category	Network/RAN
Description	Small size off-the-shelf servers
Specific solutions	Servers in data centres, enterprise IT networks
Potential benefits	Fast deployment, low energy consumption
Impact on the network node (s) (e.g. complexity)	Network node needs to be simplified/less complex in order to utilise such solutions.
Impact on the architecture	Minor - a deployment technology. Can be useful for distributing network nodes or taking functions to the edge
Maturity	Enterprise solutions available
Missing steps to achieve a good understanding of the technology	Suitability to complex telecom nodes, e.g., for L1/L2 processing.
Challenges	Management of high numbers of network nodes deployed on micro-servers, especially during distributed deployment.

N2 - Efficient/Adaptive Network Resource Usage

Technology building block name	**Traffic Optimization**
Category	RAN/ Network
Description	Adapting the transported traffic to the characteristics of the transmission path and/or the end-device using middleboxes in the network. Intelligently choosing the transmission path and last mile based on attributes of the end-device, available access technologies at the end-device's location and status of network (paths and nodes)
Specific solutions	Transcoding, transrating, caching, pre-fetching, access/path selection
Potential benefits	▪ Improved QoE (e.g., latency) for end-users. ▪ More efficient use of network resources.
Impact on the network node (s) (e.g. complexity)	Requires additional intelligence and processing capabilities in network nodes (e.g., gateways, caches). Potentially increased signalling to obtain and relay necessary information from end-devices to perform traffic optimization
Impact on the architecture	Potential definition of additional functionalities for some network nodes and new network interfaces.
Maturity	Caching and transcoding techniques are fairly mature. Pre-fetching techniques are still in experimental/early phases for mobile networks.
Missing steps to achieve a good understanding of the technology	Quantification of the benefits of different traffic optimization techniques in different mobile environments. Impact of end-to-end encryption and end-to-end content adaptation schemes (like DASH for video services) on the effectiveness of network traffic optimization techniques need to be understood.
Challenges	Application-layer (end-to-end) encryption may make it impossible to utilize some traffic optimization techniques. End-to-end content adaptation may remove the need to perform content adaptation in the network. Seamless vertical and horizontal handover among cells/hotspots and different access technologies without video service interruption are additional challenges.

Technology building block name	**Enhanced multi-operator network sharing**
Category	RAN/Network
Description	Enable sharing between two or more operators at all levels within a heterogeneous network potentially administered by multiple different organisations. Shared elements can include infrastructure, spectrum, BTS (small cells, macro, and smart base stations hosting content/services), backhaul, fronthaul, site equipment, antennas (smart antennas, massive MIMO require special consideration), core transport, platforms and sharing with non-3GPP technologies. A range of technical capabilities including spectrum sharing or reuse enhanced mobility techniques and enhanced controls for access network/access point/BTS/spectrum selection at an operator policy level. Technical capabilities should be developed to maximise the overall synergies of network sharing agreements and enable flexible business models/commercial relationships that potentially change rapidly, even on a real time basis.
Specific solutions	Steering per subscriber, per service, per radio technology, per infrastructure layer, per operator on a cost/quality optimisation basis. Has small cell to macro sharing / interoperability, BTS / backhaul / core interoperability, COMP/ICIC/coordinated scheduling, and self-optimisation capabilities across operators.
Potential benefits	▪ Retains competitive environment between MNOs by supporting network differentiation and allowing the partners to evolve at their own pace. ▪ Maintains control flexibility with the operators. ▪ Lowers cost to provide coverage and capacity. ▪ Creates perception of universal access to end user, with consistent experience across different organisations. ▪ Improves resilience, capacity utilisation and spectral efficiency. ▪ Lowers cost evolution to C-RAN and cooperative radio. ▪ Lowers power usage through higher effective density (shorter reach per site). ▪ Reduces energy by enabling the reduction of the number of radiating points outside of the busy hour.
Impact on the network node (s) (e.g. complexity)	Will increase complexity and external interfaces for most nodes in the network; has major impact on mobile device network selection mechanisms; needs new core network based platforms, such as 'access policy controller' – similar to ANDSF concept but with greater granularity; and platform to communicate service / quality / capacity available to third parties. Have design implications on Layer 1, 2 & 3 transmission/transport, multi-operator, multi-service QoS differentiation, scheduling algorithms and use of smart antenna systems. Likely to increase cross operator signalling. Increases complexity of management.
Impact on the architecture	Need for simplified and standardised inter-MNO/FNO interfaces at all levels in the network and infrastructure, such as: BTS (Macro, small cell, smart base station hosting content or services); backhaul; fronthaul; Core platforms (similar to National Roaming, MVNO scenarios) taking into account different radio technologies; OSS (service management across different organisation boundaries); BSS (enable commercial

	models). Enable knowledge of service / capacity /quality capability across operators.
Maturity	Immature Current techniques for managing inter-PLMN mobility do not support the fine-grain level of control required to give operators the needed level of flexibility and typically require bespoke design work which inhibit their application.
Missing steps to achieve a good understanding of the technology	Dedicated NGMN working group required to identify use cases and propose solutions (must be operator led, including all potential infrastructure stakeholders). Evaluating: current solutions, current and future industry models (MNO, FNO), benefits and opportunities. Cost / benefit analysis for the different scenarios.
Challenges	Must support competition and encourage investments in infrastructure. Need to agree on the ‘sharing’ models to be supported. Must enable robust and consistent services across different organisations administering different parts of the end to end service. Need to manage cross operator signalling. Current regulation and spectrum allocations management mechanisms (licensed and license-exempt) could inhibit new opportunities.

Technology building block name	**Scalable service architecture**
Category	Network/Service
Description	Ability to adapt and scale to service needs based on the use case (and mapped resource allocation)
Specific solutions	
Potential benefits	Addresses the central requirement of use-case specific service delivery and resource allocation (resource/energy/mobility on-demand) and thereby also efficiency.
Impact on the network node (s) (e.g. complexity)	Intelligent and service aware controller, with potentially centralized /distributed components.
Impact on the architecture	Granularity in functional components to be used (potentially programmed), centralized / distributed architecture
Maturity	There are trends in this direction, related to QoS, scalable media adaptation, scalable header compression innovations, best RAT selection (e.g. small cell for enterprise), and scalability in bidding and game theory in addition to awareness and cognitive environments. However, true scalable service architecture is yet to mature
Missing steps to achieve a good understanding of the technology	Potentially a combination of knowledge-based environment and end to end intelligence along with programmable networks to map component functions of re-usable resources to the needs
Challenges	Discussed above. Translation of the concept/requirement to an enabling environment

Technology building block name	**Big data**
Category	Network/Service
Description	To capture, analyse, make usable and leverage the vast amount of data available in many instances of content/service delivery. Additionally, along with behaviour, context and proximity aspects, captured (or discovered & provided) by user devices, social media/networks, content/service delivery, user-data management, research and trial data, machine/sensors (including discovery) and IoT.
Specific solutions	
Potential benefits	▪ Enhance / enrich service, performance or provisioning / management. ▪ Can potentially enhance user experience on a case by case basis, and match data to user needs.
Impact on the network node (s) (e.g. complexity)	Complexity – mechanisms to go from Big Data to usable "Small" data need to be addressed. Signalling and API exposure need security and protection measures.
Impact on the architecture	Balance between openness (capture) & control (protect). Distributed APIs with policy/security rules. Traffic optimization for prioritization, offload, etc.
Maturity	Many innovations and implementations in this area – broad term that can apply to multitude of scenarios. Yet to fully mature.
Missing steps to achieve a good understanding of the technology	Efficient and secure way to leverage vast amount of data so that benefits exceed the cost and risks. Turning Big Data into usable Small Data minimizing complexity and risks with the ability to control what and when. General resilience for massive number of nodes, devices/sensors, data, and signalling.
Challenges	See above. Also note data is often large in volume, vast in variety of types and relevance. Also there can be distinction in how fast the data is needed, creating related challenges. (e.g. real-time for proximity-based scenarios)

Technology building block name	**Context-aware/User centered network**
Category	RAN/Network
Description	Big data based behaviour aware: Network pre-judges user behaviour based on big data e.g. frequent passed trail, frequent mobility state, time specific behaviour and service preference; parameter pre-configuration according to user behaviour Service aware : Service type identification; Improve QoE by e.g. introducing finer granularity of QCI ; Reduce signalling overhead for OTT User-following: Besides ultra-dense network related solutions and 3D MIMO, UE specific virtual cell ID and RS configuration can be further studied, and UE staying in the same cell (from its own perspective) while handing over; UE specific network topology, e.g. D2D, user cooperated multi-point transmission Moving network: New network architecture for high speed scenarios, e.g. Mobile relay Mobile core network can be sliced into several overlay core networks serving different type of users, such as M2M users and so on.
Specific solutions	Parameters pre-configuration, finer granularity of QCI, UE specific virtual cell ID and RS, mobile relay, etc.
Potential benefits	Almost zero delay handover, reduced signalling overhead, edgeless network, support high speed scenario, etc.
Impact on the network node (s) (e.g. complexity)	Centralized or distributed content, context storage and analysis, dynamic reconfiguration of NW for specific UE, etc.
Impact on the architecture	Introduces mobile relay node, D2D topology, etc.
Maturity	Network cache has been deployed; mobile relay and D2D has been studied in 3GPP
Missing steps to achieve a good understanding of the technology	Although some of the research points (e.g. D2D) has been discussed and explored in telecom/mobile domains, the technology needs to be further defined, tested and proven.
Challenges	Network stored content and context analysis

Technology building block name	**Content-optimization and adaptive streaming**
Category	Network/Device/Application layer
Description	Use of client-side and server-side techniques to adapt content delivery to path characteristics and the attributes of the end-device.
Specific solutions	Dynamic adaptive streaming over HTTP (DASH), Apple HLS, Microsoft Smooth Streaming, Adobe HTTP Dynamic Streaming
Potential benefits	▪ Improved QoE for end-users through adaptation of content to characteristics of end-device and transmission path. ▪ Tailored use of radio resources depending on congestion status.
Impact on the network node (s) (e.g. complexity)	Potentially small increase in required processing capabilities in end-devices, otherwise, no impact foreseen.
Impact on the architecture	Potentially impacts the design of network traffic optimisation schemes
Maturity	Adaptive streaming techniques and content optimization have been well studied in literature and there is deployment experience in real networks. However, tuned/optimized adaptive streaming for mobile networks still needs improvement.
Missing steps to achieve a good understanding of the technology	Impact of different content optimization techniques from different content-providers on the effectiveness of adaptive streaming in different mobile environments needs to be better understood. Interaction with network traffic optimisation schemes need to be better understood.
Challenges	Convincing all content providers to deploy such technologies.

Technology building block name	**Intelligent heterogeneous management**
Category	Network
Description	The ability to manage networks in fully heterogeneous environments
Specific solutions	Technologies include NFV, SDN, SON, and Service Awareness
Potential benefits	▪ Efficient and granular management in presence of heterogeneity ▪ Cost efficiency ▪ Improved availability
Impact on the network node (s) (e.g. complexity)	Adds functions for management
Impact on the architecture	Change of architecture, e.g. decoupling control and data planes, virtual machine infrastructure, APIs, SON centralized/distributed architecture - introduces logically unified control and visibility along with potentially distributed and localized allocation of functions
Maturity	SON is well defined and further maturing. Growing trends and initiatives for virtualization and SDN not fully mature in mobile communications environment. Awareness and generally intelligent responsive networks with granular management need to mature.
Missing steps to achieve a good understanding of the technology	As discussed above. NFV and SDN used in mobile environment. SON maturity along with densification (backhaul options and multiband). Leveraging the commonality in load balancing and energy efficiency.
Challenges	Major transition. Also see above.

Technology building block name	**Embedded measurement of network performance**
Category	Network Management (OSS)
Description	To capture customer experience, device internal system measurements can be used to get and classify live quality information. Accounting for these will lower the need for external probes and avoids additional test traffic. Getting performance data from live customer traffic verifies actual instead of indicated QoE (Quality of Experience) (deduced from test traffic performance). The goal is fault isolation, to be able to tell where a problem occurs and more importantly, where a problem does NOT occur. Embedded network performance counters can help isolate elements to make investigations more efficient.
Specific solutions	The network elements themselves should have enough counters/ storage capability so that external probes are not needed and incorporate the ability to log control plane traffic. Some examples key counter functionalities: ▪ Radio report retransmits, packet delay variations (PDV) or sustained loss of forwarding (in milliseconds) of packets where packets were stored in the buffer ▪ VoIP (VoLTE for instance) codecs can report packet data behaviour after each call with such indicators as PDV buffer misses (packet didn't arrive in time), packet losses (packet was lost), consecutive packet losses (number of lost packets, max number of consecutive packets lost in a row, number of packet loss events), packet that arrived out of order, etc. ▪ Preconfigured thresholds identifying what is "normal", for instance if a packet couldn't be delivered within 50 milliseconds of storage in the buffer, it's considered abnormal radio behaviour.
Potential benefits	▪ Obtain valuable information from live customer experience ▪ Avoid additional investments on external probes ▪ Avoid additional test traffic and additional network load ▪ Enable proactive optimization of customer experience
Impact on the network node (s) (e.g. complexity)	
Impact on the architecture	Impact in the OSS area of Test & Diagnostics. Will affect probe design and handling by enforcing probe-less operations.
Maturity	Immature - even if current devices already capture some data, this is often not leveraged for measurement of customer experience
Missing steps to achieve a good understanding of the technology	Dedicated NGMN working group required to identify use cases and propose solutions (must be operator led, including all potential infrastructure stakeholders). Need to evaluate current solutions, current and future industry models, benefits and opportunities. Cost / benefit analysis for the possible scenarios
Challenges	To be determined

N3 – Other Enablers

Technology building block name	**Technologies for massive connectivity**
Category	RAN/Network
Description	There are a wide variety of small packets transmissions with different QoE (Quality of Experience) for both M2M and H2H, e.g. Periodic keep-alive packets Bursty Instant Messages Real-time critical message delivery … These small packet transmissions may cause frequent RRC transitions and contribute to network signalling congestion. Moreover, the current RRC transitions may introduce extra delay and thereby cannot satisfy the real-time requirement for some applications generating small packets transmissions. 3GPP is looking for signalling optimization of small packet transmission. Evolutional and some revolutionary mechanisms need to be devised to address this for 5G.
Specific solutions	Radio/CN signalling reduction needs to be investigated further, for example UL Data Transmission integrated with Random Access. Some new scheduling without too much signalling overhead should also be studied, for example Periodic scheduling of Keep-Alive. Contention based access with consideration to network loading and trade-off collision probability are promising candidate technologies. Furthermore, a mechanism for rapid RRC state transition needs to be devised.
Potential benefits	Significant signalling reduction, more rapid message delivery and more efficient network loading
Impact on the network node (s) (e.g. complexity)	Not a big impact on network node design. New signalling procedures, new scheduling mechanism and rapid RRC transition is required.
Impact on the architecture	No impact
Maturity	3GPP has carried out some initial study on this and more research is needed.
Missing steps to achieve a good understanding of the technology	Need to agree on what kind of small packets with specific QoE needs be optimized,
Challenges	To be explored

Technology building block name	**All optical transport network with optical router/switch**
Category	Network/Transport network
Description	Optical routing/switching minimizing opto-electric conversions in the transport plane. Electric routers/switches would be replaced by optical routing/switching devices.
Specific solutions	WDM, PON, Optical cross-connect, MEMS, All-optical switches, GMPLS, Hybrid optoelectronic switch/router, OTN, etc.
Potential benefits	High bandwidth with very low energy consumption. Have the potential to accommodate future traffic growth and reduce the high energy consumption typical in conventional routers.
Impact on the network node (s) (e.g. complexity)	Network nodes require electrical conversion prior to the execution of its functions. Therefore, placement of the network nodes can be affected by the "All Optical Transport Network" concept.
Impact on the architecture	Minor. However, if mobility can be supported at Layer 1 (λ) or Layer 2 (optical packet), impact could be major.
Maturity	Investigated in research projects. Industrial research labs reported working prototypes. Ongoing research for the affinity with SDN type transport networking.
Missing steps to achieve a good understanding of the technology	Control plane design for the optical transport plane Degree of "all Optical-ness"
Challenges	Dynamic routing, applicability to large scale telco networks

Technology building block name	**Information-centric networking**
Category	Network
Description	A topic of computing networks and "Future Internet" to migrate from a host-centric and node-centric model to a content-centric, data-oriented and information-centric networking model, with intrinsic focus on named information objects in-network caching and (name-based) routing.
Specific solutions	There have been a number of research initiatives such as Named Data Networking, Content-Centric Networking and 4WARD among many others. These have developed communication / networking models, some used by the research community for (open-source) implementations. They generally have commonalities in terms of "publish/subscribe" paradigm, caching and content-based security models, but are not all identical in detailed schemes and modelling.
Potential benefits	ICN is expected for broad and efficient distribution and manipulation of content. It is claimed to provide efficiency and scalability by focussing on the content instead of a point-to-point communication model.
Impact on the network node (s) (e.g. complexity)	The ICN architecture is "content"-based, with caching and named-based routing. There is a need for new routers and added functionalities (such as storage, logic to store content and content check procedures) for response to requests and delivery. Varying models and simulation-based implementations may include multicast functions, overlay nodes, etc.
Impact on the architecture	A different model at least at higher layer of content identification and distribution. A focus on information objects, in-network caching, receiver-driven model etc. As opposed to traditional node-centric approach. Also see impact on nodes above.
Maturity	See "solutions" above. The maturity is more or less limited to a great deal of research for several years on future Internet and focus on content distribution. Enabling over SDN transport is validated in research. The use in (mobile) communication systems is yet to be designed and subsequently mature.
Missing steps to achieve a good understanding of the technology	Lack of a detailed design or model that is applicable to a 4G / 5G environment and a clear view of migration. This also better identifies the end to end impact on nodes and architecture. The "gains" from using ICN are not understood in concrete terms; how much do we gain and how do we measure? Much of the current reporting points to simulations for a high-level comparison with a traditional Internet model.
Challenges	There is a great deal of support in research but also some concerns , e.g. privacy concerns, lack of improvement in network performance and, if it does provide greater content security and routing stability, whether these benefits only depend on a full ICN scheme in order to be realized. Naming is required and modelled, but itself can pose a challenge. Similarly, there are concerns with reliability and robustness; and as noted above, real-time processing performance and manageability. Migration, ecosystem and business models/incentives are yet to be clarified.

Technology building block name	**Privacy and Security Aspects**
Category	RAN/Network
Description	There are different aspects related to security that will play an important role for 5G design, including: ▪ Radio link encryption of user traffic. Most applications that require security often implement it themselves, for instance using TLS/SSL, IPsec or some other application-specific security. Given this, how shall we handle link encryption of user traffic in 5G? ▪ Security-design for low-latency use cases. Some 5G use cases require extremely low latency – including the latency of initiating communications. This will be an important shaping factor for the security design. ▪ Location and identity privacy will require improvements with respect to current solutions used for 4G.
Specific solutions	For radio-encryption of user traffic, it will be useful to discuss the role of radio-bearer encryption vs bearer-independent solutions. For low-latency applications, consideration should be given to ideas such as establishing shared keys between entities in anticipation that they may need to communicate. Furthermore, multiple-hop security (where intermediate nodes need to decrypt and re-encrypt data) should be avoided. Also, rather than encrypting some traffic – which requires the encryption key to be established before the first bit can be sent – transmission can potentially be started straight away with integrity protection included in the data stream, and the connection torn down if integrity protection fails For location and identity privacy, solutions for improvements with respect to the state-of-art without excessive overhead should be considered.
Potential benefits	Potential benefits include a faster handling of security procedures and better protection of location and identity privacy.
Impact on the network node (s) (e.g. complexity)	Further work needed.
Impact on the architecture	Further work needed.
Maturity	No mature solutions have been proposed to-date.
Missing steps to achieve a good understanding of the technology	The main step is to decide how to handle radio-encryption of user traffic due to the fact that an increasing fraction of application traffic terminates on the Internet and the cellular operator cannot protect the E2E internet path. In addition, the same applications run over both cellular and WiFi and the latter will often be poorly protected. While radio bearer encryption may still be worthwhile, it is also useful to consider whether the 5G security architecture can create additional business value by facilitating bearer-independent (e.g., higher layer) encryption and extending to servers on the internet, or extending to device-to-device communications
Challenges	See above

Technology building block name	**Mesh networking**
Category	Network/Transport
Description	Several radio devices cooperating to pass data between two endpoints through a series of short intermediate hops. These devices could be terminals, base stations, etc.
Specific solutions	Backhauling, fronthauling
Potential benefits and requirements addressed	• Could reduce costs to provide fronthaul and backhaul, potentially increasing the economically viable coverage range of networks • Increases reliability and redundancies through self-forming and self-healing
Impact on the network node (s) (e.g. complexity)	Requires backhauling and fronthualing interfaces and capabilities on all radio nodes.
Impact on the architecture	Big impact on the network topology.
Maturity	Standardized in IEEE 802.11 Single hop relay standardized in 3GPP Release-10
Missing steps to achieve a good understanding of the technology	
Challenges	Achieving security for the data traversing the network requires trust in all cooperating nodes, which could limit potential applications. Application to latency-critical use cases still requires some work.

Technology building block name	**Enhanced fronthauling**
Category	RAN/Transport Network
Description	Current global fronthaul standard (between BBU and RRU) is ETSI ORI (Open Radio Equipment Interface), which builds on the Common Public Radio Interface (CPRI) specification. Routing using CPRI-specific transport may lead to restrictions of flexibility in wide-scale C-RAN deployment and the baseband virtualization, and also large transport bandwidth may cause restrictions in some deployments. Also, massive MIMO, enhanced cooperation between nodes, new frequency bands and RAN sharing may lead to scalability issues.
Specific solutions	Quantized I/Q fronthaul needs to be maintained as an option because it allows for forward and backward compatibility and is technology neutral. A different functional split between BBU and RRU (i.e., with more processing at the RRU) may help minimize the fronthaul bandwidth while still supporting the full radio features. RRUs supporting this technology-specific fronthaul approach could fallback to using I/Q based fronthauling to maintain forward/ backward compatibility. Ethernet and IP packetisation can provide statistical multiplexing gain, more transport network flexibility/ harmonisation and reduce switching complexity at the BBU. All fronthaul interfaces need to be open to allow multi-vendor operation.
Potential benefits	Improved cost-efficiency and system performance, while retaining the multi-technology and future-proofing advantages of existing interface.
Impact on the network node (s)	Different BBU and RRU split seems to have quite large impact. Usage of Ethernet seems a small change.
Impact on the architecture	Fronthaul needs to be part of the overall architecture. Needs to address heterogeneous networks and be included in overall O&M/ SON framework
Maturity	ORI interface baseline already standardised for existing RATs. NGMN has conducted a study on function split between BBU and RRU for LTE, which may serve as a basis for future work; IEEE project started to study CPRI packetization within Ethernet frames. Some synchronization requirements have been reflected to ITU-T to call for more work.
Missing steps to achieve a good understanding of the technology	Need consensus on the requirements in the operator community first.
Challenges	The design of new functional split between BBU and RRU, ability of Ethernet to meet fronthaul requirements for synchronization, latency and jitter whilst maintaining commonality with other transport nodes.

Commercial Address:
ngmn ltd. • Grosser Hasenpfad 30
60598 Frankfurt • Germany

Registered Office:
ngmn ltd. • Reading Bridge House • George Street
Reading • Berkshire RG1 8LS • United Kingdom

Phone +49 69/9 07 49 98-0 • Fax +49 69/9 07 49 98-41 • office@ngmn.org • www.ngmn.org