NEXT GENERATION INTERNET

Grant Agreement No.: 957216
Call: H2020-ICT-2018-2020
Topic: ICT-56-2020
Type of action: RIA

inGenious

D3.1 Communication of IoT Devices

Revision: v.1.0

Work package	WP 3
Task	Tasks T3.1, T3.2 and T3.3
Due date	31/05/2021
Submission date	31/05/2021
Deliverable lead	SEQ
Version	1.0
Editors	Efstathios Katranaras (SEQ), Guillaume Vivier (SEQ)
Authors	Efstathios Katranaras (SEQ), Guillaume Vivier (SEQ), Ignacio J. Garcia Zuazola (NOK), Jaime Ruiz Alonso (NOK), Ignacio Benito Frontelo (NOK), Ahmad Nimr (TUD), Roberto Bomfin (TUD), Ivo Bizon Franco de Almeida (TUD), Albert Marín-Bernad (UPV), David Gomez-Barquero (UPV), Cristina Avellan (UPV), José Luis Cárcel Cerver (FV), Joan Meseguer (FV), Manuel Fuentes (5CMM), Miguel Cantero (5CMM), David Martín-Sacristán (5CMM), Carsten Weinhold (BI), Clemens Saur (NCG), Jose Costa-Requena (CMC), Luis Castillo (NED)
Reviewers	Carsten Weinhold (BI), Nuria Molner (UPV), Efstathios Katranaras (SEQ), Guillaume Vivier (SEQ), David Martín-Sacristán (5CMM), Eddy Higgins (STE), Nils Asmussen (BI)
Abstract	This document describes how iNGENIOUS plans to evolve devices and communication solutions to support Next-Generation Internet of Things (NG-IoT) scenarios. In particular, this deliverable has the objective of capturing the state of the art (SoA) technologies, from the device

	standpoint, supporting IoT and the planned innovations of iNGENIOUS in this regard, which are also mapped to the use cases of iNGENIOUS. The document covers three aspects of IoT devices, namely, i) connectivity, ii) local computation, and iii) service axis.
Keywords	Internet-of-Things, 5G, IoT Connectivity, C-IoT, Edge IoT Sensors, Immersive services

Document Revision History

Version	Date	Description of change	List of contributor(s)
V1.0	31/05/2021	EC Version	Efstathios Katranaras (SEQ), Nuria Molner (UPV)

Disclaimer

Copyright notice

Project co-funded by the European Commission in the H2020 Programme		
Nature of the deliverable:	R*	
Dissemination Level		
PU	Public, fully open, e.g. web	✓
CL	Classified, information as referred to in Commission Decision 2001/844/EC	
CO	Confidential to iNGENIOUS project and Commission Services	

** R: Document, report (excluding the periodic and final reports)*

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

OTHER: Software, technical diagram, etc.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Executive Summary

This document describes how iNGENIOUS plans to leverage and evolve devices and communication solutions to support Next-Generation Internet of Things (NG-IoT) scenarios. In particular, this deliverable has the objective of capturing the state-of-the-art (SoA) technologies, from the device standpoint, supporting IoT and the planned innovations of iNGENIOUS in this regard, which are also mapped to the use cases of iNGENIOUS. The document covers three aspects of the IoT devices, namely, connectivity, local computation, and service axis, which are briefly summarized in the following:

- Regarding connectivity, several radio access technologies are considered to support a variety of IoT devices and services with diverse requirements. Capabilities and limitations of each technology are investigated and evolution directions are explored to lower the IoT devices cost for communication by investigating new innovative solutions for air interface (leveraging improvements discussed in the standardisation bodies), flexible software defined radio, and versatile modem for 5G communication.
- In the local computation side, ultra-safe low-power dedicated platforms are considered to give to the edge IoT devices and their components ability to locally, cost-optimised and securely process high volumes of data. To this end, context-based neuromorphic computing, and isolation-by-default architectures in combination with a hardware root-of-trust are explored.
- Finally, the service axis is considered to evolve operation of IoT systems in industrial and logistic environments by enabling tele-operation driving via a new generation of immersive devices and applications. On this account, immersive devices such as head-mounted displays and haptic gloves are explored for integration into a mixed reality cockpit in order to fulfil the constrained tactile requirements of future immersive IoT use cases.

Based on the state of the art review, this document provides some elements of thoughts and ideas for research and development to further guide the technical work that will be tackled in WP3 in the next months of the project. It therefore sets the scene and guidelines for the project, with respect to device aspects.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Table of Contents

Co-funded by the Horizon 2020
Framework Programme of the European Union

List of figures

Co-funded by the Horizon 2020
Framework Programme of the European Union

List of tables

Co-funded by the Horizon 2020 Framework Programme of the European Union

Abbreviations

3GPP	3rd Generation Partnership Project
5GS	5G System
ADC	Analog-to-Digital Converter
AGV	Automated Guided Vehicle
AMF	Access and Mobility Management Function
AR	Augmented Reality
CE	Coverage Enhancement
C-IoT	Cellular IoT
DAC	Digital-to-Analog Converter
DCI	Downlink Control Information
DL	Downlink
(e)DRX	(extended) Discontinuous Reception
EC-GSM-IoT	Extended Coverage GSM in the context of IoT
eMBB	Enhanced Mobile Broadband
eMTC	Enhanced MTC
(HD/FD)-FDD	(Half-duplex/Full-duplex) Frequency Division Duplex
FPGA	Field-Programmable Gate Array
GSM	Global System for Mobile communications
HARQ(-ACK)	Hybrid Automatic Repeat Request (Acknowledgement)
HW	Hardware
IEEE	Institute of Electrical and Electronics Engineers
IIoT	Industrial IoT
IMT	International Mobile Telecommunications
iNGENIOUS	Next-Generation IoT solutions for the universal supply chain
IoT	Internet of Things
IP	Internet Protocol
ITU	International Telecommunication Union
KPI	Key Performance Indicator
LPWAN	Low-Power Wide-Area Network
LTE	Long-Term Evolution
LTE-M	LTE for MTC
M2M	Machine-to-Machine
MAC	Medium Access Control Layer
MBB	Mobile Broadband
MCS	Modulation and Coding Scheme
MEC	Multi-access Edge Computing
MIMO	Multiple-Input Multiple-Output
mMTC	Massive MTC
mmW	Millimetre Wave
MR	Mixed Reality

Co-funded by the Horizon 2020 Framework Programme of the European Union

MTC	Machine-Type Communications
NB-IoT	NarrowBand IoT
NG-IoT	Next-Generation IoT
NR	New Radio
NTN	Non-Terrestrial Networks
OS	Operating System
PA	Power Amplifier
PDSCH/PDCCH	Physical Downlink Shared/Control Channel
PHY	Physical Layer
PU	Processing Unit
PUSCH/PUCCH	Physical Uplink Shared/Control Channel
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RAN	Radio Access Network
RedCap	Reduced Capability
RF	Radio Frequency
RRC	Radio Resource Control
Rx	Receive
SA/NSA	Stand Alone / Non-Stand Alone
SDR	Software Defined Radio
SoA	State of the Art
SW	Software
TCU	Trusted Communication Unit
TDD	Time Division Duplex
ToD	Tele-operation Driving
TSN	Time Sensitive Networking
Tx	Transmit
UC	Use Case
UE	User Equipment
UL	Uplink
UPF	User Plane Function
URLLC	Ultra Reliable Low Latency Communications
VR	Virtual Reality
XR	Extended Reality

Co-funded by the Horizon 2020 Framework Programme of the European Union

1 Introduction

1.1 Objective of this Deliverable

This deliverable is the first report of the work package 3 (WP3) of the iNGENIOUS project. The report aims primarily at capturing the current status (at end of 2020, early 2021) of technologies to support Internet of Things (IoT), from the device standpoint, in the specific context of iNGENIOUS, as recalled in Section 1.2. Starting from this state-of-the-art (SoA) survey, the report also identifies directions for improvement for the device and also from the communication system perspective. These directions are setting the baseline for the research and technical activities to be undergone within WP3.

In iNGENIOUS, several next generation IoT use case (UC) scenarios are envisaged, each with different devices and communication requirements. For that reason, deliverable D3.1 explores several technologies from the air interface point of view as well as the respective devices' capabilities in order to address those requirements. While the legacy technologies are scrutinized, the document also focuses on ongoing evolution from 3rd Generation Partnership Project (3GPP), the cellular standards body, towards the support of new generation 5G IoT scenarios and opportunities for technology improvement from WP3 research and technical activities. In addition, this document describes how iNGENIOUS plans to evolve the hardware and software architectures of the various components of the IoT devices in order to achieve cost- optimized use case specific system solutions (via edge sensors with context-based neuromorphic edge clustering) as well as give the device the ability to process data locally, securely, and at low-power (via tile-based hardware/software architecture). Finally, the document provides an initial description of the immersive devices (such as Head-Mounted Displays and Haptic Gloves) that will be developed to improve operation of IoT systems in industrial and logistic environments.

In summary, this document serves as a benchmark report and describes the planned work of iNGENIOUS for innovation within *functionalities* related to three general aspects of IoT devices: i) connectivity, ii) local computation, and iii) service axis. These functionalities are part of the Device and User Equipment (UE) domains as captured in iNGENIOUS network architecture provided in the deliverable 2.1 (D2.1) of iNGENIOUS [1] and partly reproduced in Figure 1.1 below with focus on these WP3-relevant aspects.

Co-funded by the Horizon 2020
Framework Programme of the European Union

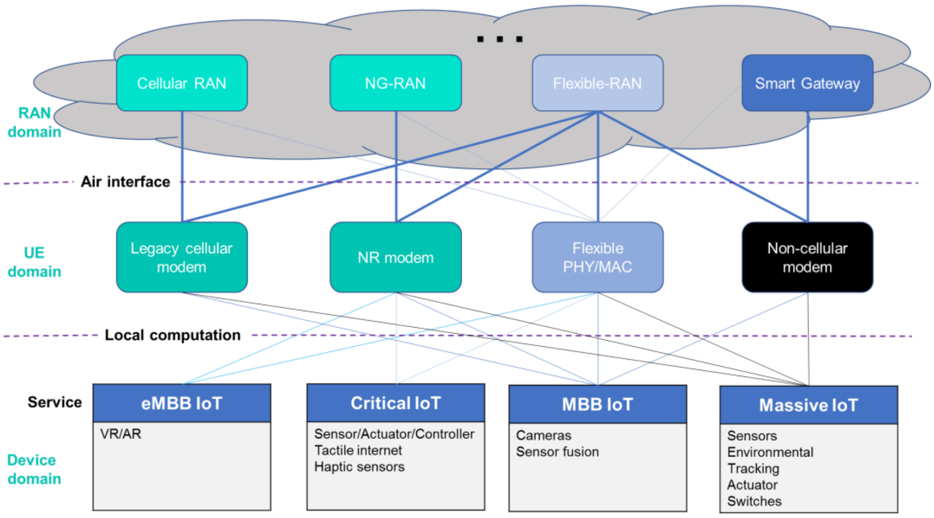


Figure 1.1. Device and UE domains and respective functionalities within iNGENIOUS network architecture

1.2 Role of WP3 in iNGENIOUS

Industries and other verticals are experiencing the benefits of digitalization thanks to the integration of wireless connectivity solutions. The predicted increase of connected devices anticipates the need to connect everyone and everything, enabling the next generation of IoT.

The term IoT describes the interconnection between any devices or machines over the Internet to enable a seamless communication with the minimal human interaction. The lowest level within the IoT system architecture is the *IoT device layer*. This essentially includes the “things” that interact with the physical world. The purpose of IoT devices is wide open, ranging from simple devices (e.g. sensors, wearables, asset trackers, security systems, etc.) to sophisticated industrial robots, automation tools and connected vehicles.

In WP3, iNGENIOUS aims to evolve the hardware and software architectures of IoT devices as well as their communication. The goal is to address the current limitations and ease adoption of devices in next generation IoT scenarios. Figure 1.2 depicts the UC scenarios of iNGENIOUS with various IoT devices, which are described in detail in D2.1 [1].

Co-funded by the Horizon 2020 Framework Programme of the European Union

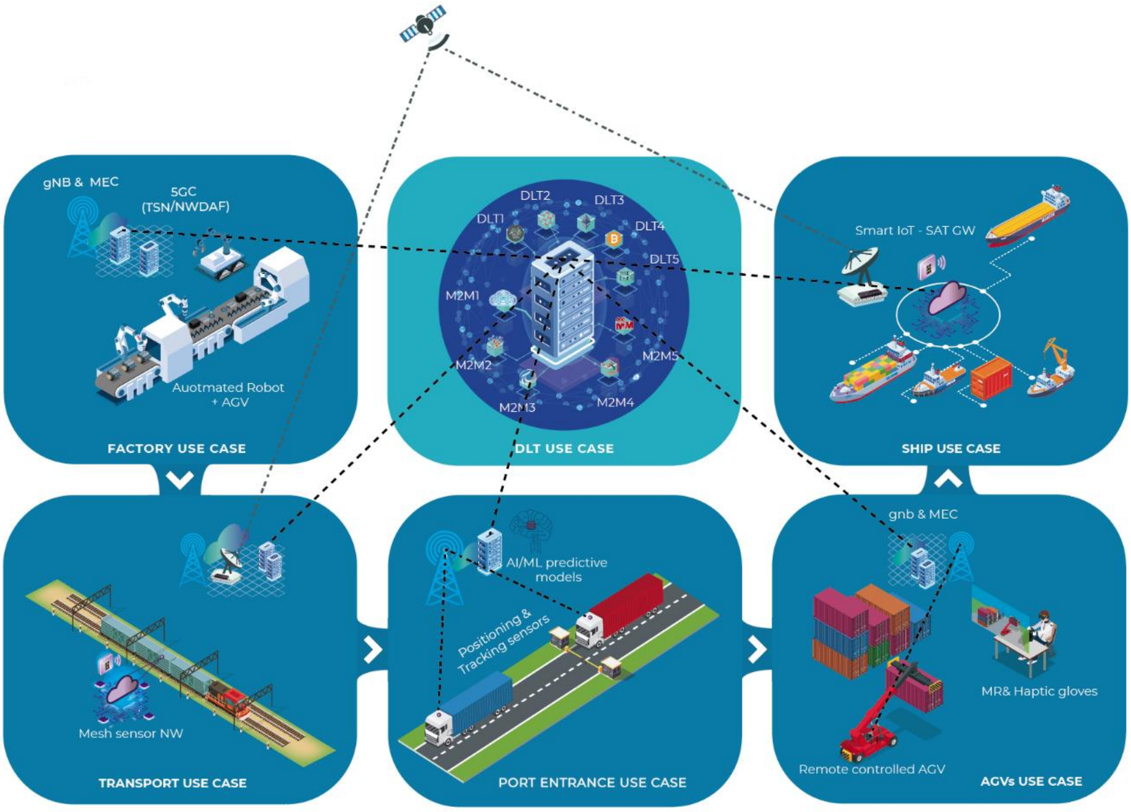


Figure 1.2. iNGENIOUS next generation supply chain use cases

Three main aspects of evolution are considered in WP3:

- **Connectivity**

  To lower the IoT devices cost for communication, in terms of computational complexity, power consumption, latency, flexibility, etc., by investigating new innovative solutions for air interface (leveraging improvements discussed in the standardisation bodies), flexible software defined radio (SDR), and versatile modem for 5G communication.

- **Local computation**

  To give to the edge IoT devices and their components ability to locally, cost-optimised and securely process high volumes of data, with context-based neuromorphic computing, joint componentised micro-kernel operating system (OS), and tile-based hardware architecture with embedded artificial intelligence (AI) capabilities.

- **Services**

  To enable immersive applications via head-mounted displays and haptic devices in order to fulfil the constrained tactile requirements of future immersive IoT use cases.

These three aspects are respectively addressed in the three defined tasks of WP3. Each task will consider the requirements defined in WP2 as the basis for the research activity and coordinate closely with WP4 which focuses on radio access network (RAN) and core network (CN) aspects. The outcome of

Co-funded by the Horizon 2020 Framework Programme of the European Union

the work will be used in prototypes to be integrated and validated into the WP6 planned trials as well as in research proposals to be submitted to standardization bodies as part of WP7 efforts.

Accordingly, a set of objectives has been defined in WP3, namely:

- to define and promote **evolutions of IoT connectivity** in line with standardization framework as well as to investigate innovative solutions to support the core and access technologies adopted within iNGENIOUS,
- to define the **forthcoming hardware (HW) and software (SW) architectures** for immersive, tactile and more secure IoT devices, and
- to **develop proof-of-concept** (PoC) of the hardware and software architectures.

1.3 Role of WP3 in the Use Cases

The goal of this section is to describe how WP3 is related to UCs depicted in Figure 1.2 and defined in deliverable D2.1 [1]. The six UCs together have the objective of demonstrating how the technical innovations of iNGENIOUS will enable digitalization and monitoring of the supply chain ecosystem as a whole, enhancing factory operations, transportation and maritime port operation. In the following, the six UCs and their relations to the explored functionalities within WP3 (for connectivity, local computation, and immersive services) are briefly described. A more detailed description of the UCs is found in D2.1 [1].

1.3.1 Automated robots with heterogeneous networks (Factory UC)

Factory UC considers automated robots in smart factories that make use of the tactile internet. In this scenario, the sensors and machinery synchronously work with latencies of few milliseconds, where the interoperability with wired Time Sensitive Networking (TSN) environments is also explored. This UC will enable automated robot and automated guided vehicle (AGV) control in industrial environments based on the design of a smart distributed application that will leverage different types of sensors, actuators and parallel control loops for connecting machines and humans.

The contributions of WP3 to this UC include the use of various SoA IoT connectivity solutions to interconnect heterogeneous IoT and tactile devices and realise and test industrial IoT with low-latency and ultra-reliability requirements as well as its integration with massive and broadband IoT. Non-cellular-based (Section 2.1) and cellular-based (Section 2.2) solutions will be used for such connectivity. The 5G connectivity for AGV will be supported with a versatile 5G modem (Section 2.4.3). Potentially, reduced capability 5G devices (Section 2.2.2.2) may be used for demonstration along with supported evaluations from respective link/system-level simulations (Section 2.4.1). In addition, software defined approaches for flexible hardware/software architecture (Section 2.4.2) will be used to develop a real-time waveforms framework to support a smart air interface and flexible resource allocation.

Co-funded by the Horizon 2020 Framework Programme of the European Union

1.3.2 Transportation platforms health monitoring (Transport UC)

Transport UC has the objective of showing that asset health tracking can lead to low operational costs and high asset availability. It uses new data-based services provided by low-power edge distributed networks and intelligent sensor modules installed in the transportation platforms. Specifically, this UC will enable the monitoring of health for transportation platforms with a particular focus on railway transportation. The monitoring of health conditions will reside in the design of neuromorphic sensors able to gather and process data on the edge of the network while bringing low cost and power and high life expectancy. For enabling the exchange of data between smart edge sensors and platforms, near continuous connectivity to the edge will be enabled by exploiting terrestrial and non-terrestrial access networks.

The contributions of WP3 to this UC include a sensor platform with diverse IoT connectivity where neuromorphic context-based data clustering at edge (Section 3.4.1) will be developed to demonstrate edge sensing applications for dedicated and exploratory data driven edge tasks. Situation-based energy optimized edge computing (Section 3.4.2) for condition and novelty monitoring by edge sensors will also be investigated. For IoT connectivity, the use of low-power legacy cellular-based IoT (Section 2.2.1), to be integrated with sensor platform, is to be demonstrated, while other SoA connectivity solutions (Section 2.1) will also be considered. In addition, satellite connectivity for remote underdeveloped regions is considered to be demonstrated in this UC, therefore, the innovative concepts to be investigated for satellite-based communication and respective link/system-level simulations will be applicable (Section 2.4.1). Furthermore, an isolation-by-default computer architecture (Section 3.4.3) will be demonstrated via the tile-based hardware/software co-designed platform using a microkernel-based OS to secure sensor access and the communication channel. A minimal root of trust, integrated into the hardware, and corresponding OS support for remote attestation and secure software updates in the platform will also be developed (Section 3.4.4).

1.3.3 Situational understanding and predictive models in smart logistics scenarios (Port Entrance UC)

Port Entrance UC targets the development of artificial intelligence (AI) and machine learning (ML) based predictive models to estimate and optimise truck turnaround times. By means of collecting, aggregating, and processing data using ML, this UC aims at reducing the time vehicles spend inside the port and terminal facilities as well as wait times at the port accesses. This leads to corresponding savings on direct costs for carriers, thanks to enhanced situational understanding of events in maritime ports and terminals.

The contributions of WP3 to this UC will be limited to the use of cellular SoA IoT communication devices and connectivity solutions. In particular, this UC will leverage IoT tracking devices with legacy LTE-M and/or traditional LTE connectivity (see Section 2.2.1) for obtaining the real-time position of trucks

Co-funded by the Horizon 2020 Framework Programme of the European Union

inside the port facilities. The positioning of trucks will be used for validating the estimations performed by the predictive models when calculating truck turnaround times.

1.3.4 Improved driver's safety with mixed reality and haptic solutions (AGV UC)

AGV UC aims at improving the safety conditions of workers in maritime ports and terminals, by keeping the employees away from hazardous environments such as fuel port terminals using innovative wireless networks. The idea is that the workers operate AGVs remotely using 5G wireless communications and Mixed Reality (MR) and haptic solutions, which will be integrated in an immersive remote indoor cockpit. 5G will ensure high throughput and low latency connectivity with the cockpit, while MR and haptic solutions will provide an immersive experience to the worker.

The contributions of WP3 to this UC include an immersive remote indoor MR cockpit (Section 4.3), supported by low-latency video cameras and proximity sensors installed in AGVs, which will be wirelessly connected to a 5G RAN via a versatile 5G modem (Section 2.4.3) using SoA broadband NR-based IoT air interface technology (see Section 2.2.2). Additional legacy cellular and non-cellular IoT connectivity solutions will be considered while simulations might also complement the connectivity evaluation prior experimental results. Furthermore, the tele-operation driving will be achieved by the use of immersive devices such as head-mounted devices and haptic gloves (Section 4.3.2), enabling operators to remotely control the AGVs from a safer position than outdoors where hazardous working environments and adverse weather conditions can be encountered.

1.3.5 Inter-modal asset tracking via IoT and satellite (Ship UC)

The goal of Ship UC is to provide end-to-end intermodal asset tracking with IoT and satellite connectivity for enabling enhanced real-time monitoring of shipping containers when they are transported in both terrestrial and maritime segments. In order to achieve this objective, this UC will explore communications via satellite backhaul and IoT terrestrial infrastructure, enabling real-time monitoring of cargo parameters when containers are sailing on the sea and when they approach the coast. To enable the ubiquitous coverage, sensors and IoT tracking devices will be installed on the shipping containers transported by ships and trucks on both segments, whose connectivity will be provided by a smart IoT gateway (GW).

The contributions of WP3 to this UC include the use of SoA IoT communication modems (LoRa, LTE-based IoT, NR-based IoT) and respective air interfaces technologies (see Sections 2.1.3, 2.2.1, and 2.2.2). These provide network connectivity to the heterogeneous IoT devices monitoring the container, directly to the terrestrial access network, or to the smart IoT GW installed in ship or truck. Since satellite backhaul is envisaged to be in place when the ship is travelling on the sea, innovative concepts investigated for direct access satellite-based communication and respective link/system-level simulations will also be applicable here (Section 2.4.1).

Co-funded by the Horizon 2020 Framework Programme of the European Union

1.3.6 Supply chain ecosystem integration (DLT UC)

In DLT UC, the project aims at enabling interoperability between existing machine-to-machine (M2M) communication and distributed ledger technology (DLT) solutions through the development of M2M virtualization and DLT layers. The idea is to provide two different interoperable layers in order to abstract the complexity of the underlying M2M platforms and DLT solutions, guaranteeing at the same time data privacy and security by means of encoding and anonymization techniques. These layers will be capable of securely and semantically exchange the information flows between the different actors that can take part along the supply chain ecosystem. This could be used in different segments of the supply chain such as manufacturing, transportation or logistics.

Most of the technical contribution in this UC will come from WP5 as stated in D2.1 [1]. WP3 contribution will be limited to the use of SoA communication devices and air interface connectivity solutions (Sections 2.1 and 2.2) for the set of different M2M platforms that will be used to feed data to the data virtualization layer.

1.4 Structure of the Document

This document is organized in three main chapters, each including primarily the relevant SoA in detail and, subsequently, a description of the planned WP3-related innovations as well as the present vision on how the various functionalities map to the different iNGENIOUS UCs in terms of innovation maturity. The three chapters are:

- **Chapter 2**, *IoT Connectivity*: this chapter describes the various IoT connectivity technologies that iNGENIOUS will leverage and evolve for improving IoT device cost for communication, including cellular and non-cellular air interface solutions, and respective communication devices, as well as flexible software defined radio.
- **Chapter 3**, *Ultra-safe Low-Power Dedicated Platforms*: this chapter describes the hardware and software architectures of the various components of the IoT devices that iNGENIOUS will explore and evolve to optimise power consumption as well as to harden the devices against malicious attacks to keep the IoT as a whole trustworthy.
- **Chapter 4**, *Immersive Devices and Applications*: this chapter describes the new generation immersive applications that iNGENIOUS will implement, as well as the respective immersive devices to be developed and integrated, which are key enablers for improving operation of IoT systems in industrial and logistic environments.

Co-funded by the Horizon 2020 Framework Programme of the European Union

2 IoT Connectivity

Nowadays, the communication system becomes an unavoidable aspect of any industrial or societal process. Moreover, communication systems are very diverse, from fixed to wireless solutions, with many standards or proprietary systems. In this context, it is vital to be able to determine the most appropriate communication strategy suited to the particular needs of a given business and make it successful.

A very important part of the communication system includes the radio interface technology (RIT), or also known as air interface, or radio access, and the respective communication devices, in the sense of e.g. chip, modem, or module, needed to realize the IoT device connectivity with the network entities. Operators and enterprises are looking to identify, assess, and deploy the optimum form of connectivity that meets the best quality of service for their application at the most compelling total cost, primarily including coverage and cost (for both implementation and maintenance). Selecting the wrong radio strategy can have a negative impact on the quality of service, increase costs and deteriorate security.

The connectivity requirements of IoT networks may vary depending on the implementation conditions and their usage scenario. The most relevant Key Performance Indicators (KPIs) in this context are as follows:

- **Coverage**: aimed at supporting not only short range communications of a few meters, but also long range coverage of a few kilometres in urban areas and over 10 km in rural settings.
- **Reliability**: measuring the capability of transmitting a given amount of traffic within predetermined time duration with high success probability.
- **Latency**: may have relaxed constraints in applications like smart homes, but stringent requirements in real-time or decision-making applications.
- **Throughput**: ranging from a few kbps to send status data to several Mbps (or even Gbps) to carry video information.
- **Battery life**: aimed to be long-lasting based on the RIT capabilities to minimize the power consumed by the device modem in relation to the traffic characteristics.
- **Device cost/complexity**: aimed at enabling an IoT cost-efficient network with a large number of low-complexity devices.

Various connectivity options exist currently in market to address the diverse scenario requirements. Taking a view on IoT market segmentation (see also Figure 2.4 and discussion later in the chapter) one can distinguish the main following categories based on communication devices requirements and capabilities:

- **Massive IoT**, addressing meters, sensors, trackers, wearables, etc., with infrequent, small data volumes, which requires low cost (usually narrow bandwidth) devices of extreme coverage and that may rely solely on battery.

- **Broadband (or MBB/eMBB) IoT**, addressing vehicles, gadgets, cameras, actuators, etc. with high or very high data rates (maybe at the Gbps level) and volumes, and UL heavy and/or periodic traffic, which requires device capabilities of low latency but also extended battery life, coverage and position compared to the high mobile broadband service-oriented devices.
- **Critical (or URLLC) IoT**, addressing Augmented or Virtual reality (AR/VR), mobile robots, real-time human-machine collaboration, etc., with ultra-low latency as well as ultra-reliable data delivery (within bounded latency) capabilities. A respective branch of this category can also be considered to be the *Industrial IoT* for parts of industrial system (e.g. construction, ports, etc.) that require real-time advanced automation and seamless integration into wired industrial Infrastructure.

Since not all connectivity technologies are capable to fulfil the same requirements, one IoT application may resort to one type of solution or another, depending on the capabilities of such technologies. For instance, the Low Power Wide Area (LPWA) applications are characterised by targeting long range, low-power and low-cost use cases with relaxed throughput requirements.

Motivated by the idea of pursuing a "digital society" with a myriad of interconnected devices, the radio access solutions for IoT are classified between two mainstream approaches: the so-called *cellular IoT* (C-IoT) technologies and the *non-cellular IoT* technologies. Specifically for the C-IoT, 3GPP has been developing different licensed technologies throughout the Long-Term Evolution (LTE – the 4G-oriented air interface) standardisation era, such as LTE for machine-type communications (LTE-M) and narrowband IoT (NB-IoT), targeting LPWA use cases with the respective Cat-M and Cat-NB communication devices, and more recently respective feature sets based on New Radio (NR - the new 5G-oriented air interface) like NR Industrial IoT (IIoT), Reduced Capability (RedCap) NR devices, and IoT over Non-Terrestrial Networks (NTN). Note that these technologies developed by 3GPP can also be deployed in unlicensed spectrum, leveraging the work done for instance in context of LTE and NR for Unlicensed spectrum (LTE-U and NR-U, respectively).

Until a few years ago, radio connectivity options were focused primarily on cellular connectivity, Wi-Fi connectivity or customized/proprietary private-network connectivity. Today's connectivity options however are able to take advantage of licensed and license-exempt spectrums to include C-IoT, private LTE or NR, Wi-Fi, and Low Power Wide Area Networks (LPWAN) such as LoRaWAN.

iNGENIOUS, within its WP3 task to explore IoT connectivity towards contribution to the Next-Generation IoT (NG-IoT) connectivity design, will first survey the state-of-the-art of various connectivity technologies and respective communication devices in order to identify what can be leveraged for its selected use cases and limitations of existing solutions,

Co-funded by the Horizon 2020 Framework Programme of the European Union

while the ongoing 3GPP work will also be of great focus where opportunities for innovation will be part of the planned technical research. Furthermore, NG-IoT connectivity design in iNGENIOUS targets to cover various cases that may require multiple radio solutions. To this end, it is of great interest to investigate SDR-based solution to provide a single HW/SW framework that could be applicable to multiple radio system as well as versatile communication devices that will provide functionalities not available with current IoT solutions.

In the following of this chapter, we first introduce and describe the state-of-the-art for the diverse IoT connectivity solutions that iNGENIOUS will leverage and/or target to evolve for its selected use cases. More specifically, we briefly describe in Section 2.1 the non-cellular solutions which iNGENIOUS plans to use as part of its communication technologies portfolio in its UCs. Then, in Section 2.2, we dive in detail on state-of-the-art of cellular solution with additional special focus on ongoing 3GPP work. Section 2.3 explains the aspect of baseband signal design at the transmitter and the concept of flexible SDR. Finally, Section 2.4 discusses the innovations and improvements over the state of the art that iNGENIOUS partners currently investigate.

2.1 Non-Cellular IoT

In this section, the non-cellular IoT technologies which will be considered to be leveraged in iNGENIOUS use cases are described, including Zigbee, Bluetooth, LoRa, and Sigfox. It is of course worth mentioning that there are many other “akin” systems, e.g., Eltres (which is introduced in D4.1 [3] and is under consideration in iNGENIOUS), 6lowpan, Thread, Zwave, all mesh solutions, etc. Another recent interesting technology is DECT2020 NR, standardized by the European Telecommunications Standards Institute (ETSI), which could be used to support wide range mesh networks in use cases such as logistics and asset tracking, industry 4.0 and building automation as well as condition monitoring [4].

2.1.1 Zigbee

Zigbee is a wireless technology developed by the Zigbee Alliance in 2003 as an open global communications standard to address the need of low-cost and low-power wireless IoT networks [5]. The standard was designed to rely on the Institute of Electrical and Electronics Engineers (IEEE) 802.15.4 physical radio specification, physical (PHY) and medium access control (MAC) layers, which is required for enabling wireless personal area networks (WPAN) through the use of low power radio-enabled devices. Thanks to this dependency, Zigbee devices are able to operate at one of three unlicensed spectrum bands: 868 MHz in Europe, 915 MHz in North America, and 2.4 GHz worldwide. In this spectrum distribution, 868 MHz band is split in one single channel, 915 MHz band is split in 10 channels with 2 MHz separation between 902 MHz to 928 MHz, and 2.4 GHz is composed of 16 channels with 5 MHz spacing.

In PHY layer, coding is based on the Direct Sequence Spread Spectrum (DSSS) technique and modulation on Offset quadrature phase-shift keying (QPSK). Regarding MAC layer, the basic channel access mode is carrier-sense multiple access with collision avoidance (CSMA/CA), which performs a clear

Co-funded by the Horizon 2020 Framework Programme of the European Union

channel assessment (CCA) for avoiding collisions in transmissions. Nevertheless, CSMA/CA is not used in three notable exceptions: (i) message acknowledgments, (ii) beacons that are sent on a fixed-timing schedule, (iii) devices in beacon-enabled networks that have low-latency real-time requirements since they may also require the use of guaranteed time slots. In these cases, a beacon-enabled protocol is used to dedicate a specific time slot to a particular device. The use of beacon-based techniques is especially relevant for enabling low power consumption since nodes only need to be active while a beacon is transmitted and not continuously, as for CSMA/CA techniques.

At device level, Zigbee defines three different types of devices:

- **Coordinator**: This device starts the network by selecting the channel and the personal area network (PAN) ID, buffers wireless data packets for sleeping end device children and manages other functions that define the network, secure it, and keep it healthy. Zigbee networks only have one coordinator device, which must be powered on all the time.
- **Router**: A router is a full-featured Zigbee node that can join existing networks and send, receive, and route information. Routers can buffer wireless data packets for sleeping end device children and allow other routers and end devices to join the network. There may be multiple router devices in a network and as for coordinators, routers cannot sleep, i.e. must be powered on all the time.
- **End device**: These devices can join existing networks and send and receive information, but cannot act as messenger or router between any other devices. End devices use cheaper hardware and can power itself down intermittently, saving energy by temporarily entering a non-responsive sleep mode. These devices always need a router or the coordinator to be its parent device and help them to join the network, and store messages for them when they are asleep. Zigbee networks may have any number of end devices.

Regarding network architecture, Zigbee supports star, tree, and mesh topologies. In a star topology, the network is controlled by one single device called the Zigbee coordinator. While the coordinator is responsible for initiating and maintaining the devices on the network, end devices directly communicate with the Zigbee coordinator. In mesh and tree topologies, the Zigbee coordinator is responsible for starting the network and for choosing certain key network parameters, but the network may be extend its coverage through the use of Zigbee routers. In tree networks, routers move data and control messages through the network using a hierarchical routing strategy. Tree networks may employ beacon-oriented communication. Mesh networks allow full peer-to-peer communication. Zigbee routers in mesh networks do not currently emit regular IEEE 802.15.4 beacons. Figure 2.1 shows a generic Zigbee network topology where the three types of devices are interconnected.

Co-funded by the Horizon 2020 Framework Programme of the European Union

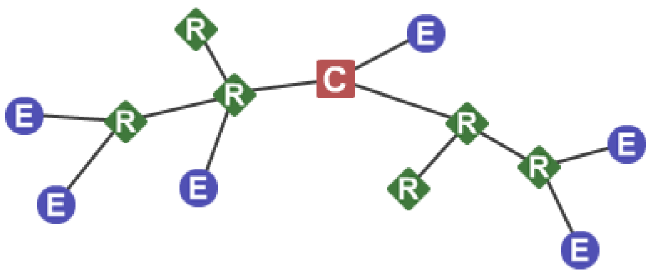


Figure 2.1. Zigbee network architecture with Coordinator, Router and End-Devices [2]

Thanks to these features, Zigbee provides energy-efficiency (several years of battery life), low cost, low data-rates (20–40 kbps at 868/915 MHz and 250 kbps at 2.4 GHz) with 2 MHz bandwidth, and short range coverage (10-300 m) capabilities. These features, combined with the decentralized and mesh network-based architecture, make Zigbee a suitable technology for wireless sensor network (WSN) based applications in agricultural and farming domains (such as irrigation management, pesticide and fertilizer control, and water quality management) where small and periodic information updates are required.

2.1.2 Bluetooth

Bluetooth (BT) is a short-range wireless technology standard designed in 1994 for exchanging data over short-range distances from fixed and mobile devices. The basic conception is to create personal area networks – typically with coverages up to 100m – with high levels of security based on ad-hoc (i.e. direct communication) technology, which allows communicating data wirelessly between paired devices [6].

Along time, Bluetooth has been upgraded through different versions that integrate improvements in speed, range and data capacity. Bluetooth technology is based on IEEE 802.15.1 standard and all the versions are compatible with their previous versions, enabling device interoperability. The list of Bluetooth standards evolution and their associated features for respective IoT devices is the following:

- **Bluetooth v1.0**: First standard of Bluetooth technology that faced many difficulties with interoperability as manufacturers struggled to make their products interoperable. It defined the baseline Bluetooth hardware for the upcoming versions.
- **Bluetooth v1.1**: Fixed the problems faced in v1.0 and added non-encrypted channels and signal strength indicators, leading to the definition of IEEE standard 802.15.1 in 2002.
- **Bluetooth v1.2**: Faster transmission speed compared to v1.1, up to 721 kbps, involving retransmission of corrupted data packets.
- **Bluetooth v2.0+EDR**: Introduced Adaptive Frequency Hopping (AFH) and Enhanced Data Rate (EDR) technology, which enables faster transmission of data allowing devices to boost maximum data transfer rate up to 2.1 Mbps. EDR, which can also provide a lower power consumption, is stated as an optional feature in this version.

Co-funded by the Horizon 2020 Framework Programme of the European Union

- **Bluetooth v2.1+EDR**: Improved the pairing experience for Bluetooth devices allowing the pairing to happen much faster and more easily thanks to a simplification of the pairing encryption, while increasing the overall security.

- **Bluetooth v3.0+ High Speed (HS)**: Allows transferring large amounts of data, reaching data transfer speeds of up to 24 Mbps thanks to the use of 802.11 protocol features.

- **Bluetooth v4.0 LE** or **Bluetooth Low Energy (BLE)**: Keeps high speed data rates while lowering the power consumption thanks to the transmission of small data packages with simple modulation. BLE was designed for enabling frequent data transmissions in smart devices such as fitness bands, smart watches, etc., leading to lower power consumption.

- **Bluetooth v5.0 LE**: BLE 5.0 is considered as a candidate technology for enabling Industrial IoT use cases related to monitoring or massive sensorisation in factories or automotive scenarios thanks to the introduction of several capabilities such as: increased broadcasting capacity, throughputs up to 2 Mbps, use of Bluetooth beacons, larger message capacity and support for IoT devices.

Bluetooth operates in the globally unlicensed Industrial, Scientific and Medical (ISM) 2.4 GHz radio frequency band, at frequencies between 2.402 and 2.480 GHz, or 2.400 and 2.4835 GHz, including guard bands (2 MHz wide at the bottom and 3.5 MHz wide at the top). In this spectrum portion, first Bluetooth versions included 79 designated Bluetooth channels with a bandwidth of 1 MHz per channel. At the latest Bluetooth Low Energy release, 40 channels of 2 MHz bandwidth were accommodated. In particular, the 40 channels are divided into three advertising channels (37, 38, and 39), and 37 data channels (0-36). While advertising channels are used for device discovery, connection establishment and broadcast transmissions purposes, data channels are used for exchanging data through bi-directional communication between connected devices.

Regarding air interface, in PHY layer, originally Gaussian frequency-shift keying (GFSK) modulation was the only modulation scheme available. Since the introduction of Bluetooth 2.0+EDR, π/4-DQPSK (differential QPSK) and 8-DPSK modulation may also be used between compatible devices. Devices functioning with GFSK are said to be operating in basic rate (BR) mode, where an instantaneous bit rate of 1 Mbit/s is possible. The term EDR is essentially used to describe π/4-DPSK and 8-DPSK schemes, each giving 2 and 3 Mbit/s respectively. In MAC layer, Bluetooth uses a radio technology called frequency-hopping spread spectrum where data is divided into packets and transmitted in each frequency channel. Bluetooth performs up to 1600 hops per second, with adaptive frequency-hopping (AFH) mode enabled.

2.1.3 LoRa

Co-funded by the Horizon 2020
Framework Programme of the European Union

LoRa is another technology operating in unlicensed spectrum, specified by an industry alliance – the LoRa Alliance – targeting to provide long-range connectivity to battery operated devices. A network employing this technology, the so called LoRaWAN, can be deployed on top of an existing Wi-Fi or 5G network for complementation and operational cost optimization. LoRaWAN requires about 2x the transmission power of Bluetooth and 5x less energy than NB-IoT with superior signal range at the cost of lower transmission rates (see Section 2.2.3).

LoRaWAN is ideally suited for Mesh networks and can be combined with various technologies in hybrid networks.

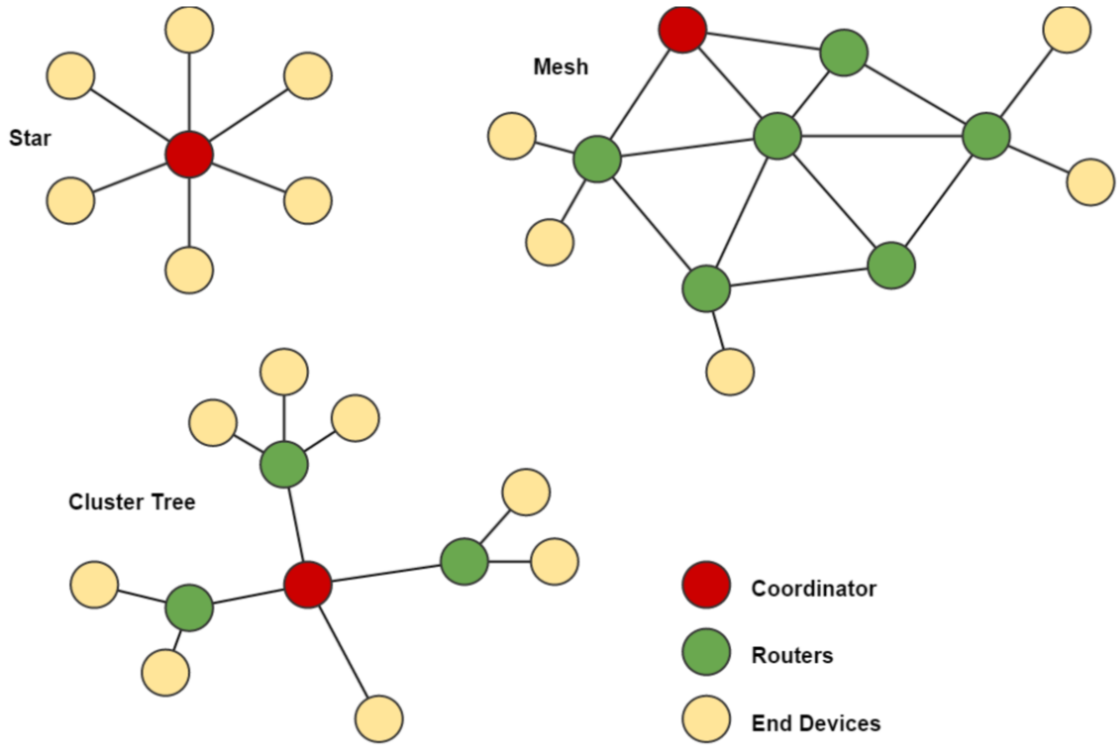


Figure 2.2. LoRa network architecture topologies

Typically, LoRa access requires a dedicated service provider such as *The Things Network*, a global community building an open-source and decentralized LoRaWAN network [7]. Recently a new innovative approach called the *Helium Network* has been deployed to disrupt the concept of dedicated service providers. Helium is a decentralized peer-to-peer wireless which builds on "public" LoRa Mesh Networks and gets financed via transmission micro payments [8]. This blockchain approach transforms IoT data transmissions into Business "Information Order Transmissions". This pay per use concept is extremely attractive for irregular push and pull information systems.

Further information on LoRa physical layer solution, spectrum usage and regulatory requirements can be found in D4.1 of iNGENIOUS [3].

LoRaWAN end-nodes are designed to serve different applications and meet different battery life and latency requirements. The three devices classes generally considered (Class A, B, and C) are also described in D4.1 [3].

2.1.4 Sigfox

Sigfox is a French global network operator founded in 2010 that builds wireless networks to connect low-power objects, such as electricity meters and smartwatches, which need to be continuously on and emitting small amounts of data. Performance-wise it is similar to LoRaWAN. More information on Sigfox technology specific details, communication stack and protocol can be found in D4.1 [3].

It is worth noting that the Sigfox business model takes a top-down approach. The company owns all of its technology, from the backend data and cloud server to the endpoints software. But the differentiator is that Sigfox is essentially an open market for the endpoints. Sigfox gives away its endpoint technology to whatever silicon manufacturer or vendor wants it as long as certain business terms are agreed upon. Large manufacturers like STMicroelectronics, Atmel, and Texas Instruments make Sigfox radios. Sigfox thinks that keeping the application cost low is the way to drive people to its market. This is different to the LoRa Alliance strategy where the specification that governs how the network is managed is relatively open. One can download the specifications and join the LoRa Alliance, and any hardware or gateway manufacturer can build a module or gateway that conforms with LoRa specifications. The "restriction" is that the only company that makes the radio for LoRa is Semtech while some other manufacturers make system-in-package devices with Semtech silicon or intellectual property inside. Thus, while the LoRa ecosystem itself is open, it does have a closed element.

2.2 Cellular IoT

Specifications and enhancements from 3GPP to support machine-type, or M2M, communications (MTC), and the resulting IoT applications, have a long history. The Extended Coverage Global System for Mobile communications in the context of IoT (EC-GSM-IoT) was firstly introduced by the 3GPP in Release 13 as an LPWA technology based on enhanced general packet radio services (EGPRS). In comparison with GSM/GPRS devices, EC-GSM-IoT was designed as long range, long battery life and low complexity system able to coexist with the existing mobile networks. It has global presence making use of four GSM frequency bands (850, 900, 1800 and 1900 MHz). It also targeted to add improved coverage by 20 dB over EGPRS, LTE-grade security, and power efficient operation [5]. Further, it could support a huge number of devices (over 50.000 per cell) and include user identity confidentiality, entity authentication, confidentiality, data integrity, and mobile equipment identification. However, it is worth noting that EC-GSM has never really been deployed; 2G MTC, as of today, operates in practice using regular GSM-based technology. While 3G is shutting down in many European countries like Germany [10], the Netherlands [11], or Denmark [12], 2G is still dominant in many scenarios. This is the case of some typical M2M applications in urban areas and, especially, sensor networks deployment in rural environments.

Co-funded by the Horizon 2020 Framework Programme of the European Union

The release of LTE made 4G competitive in the field of MTC. 4G capabilities are far more powerful than the precedent cellular networks in terms of throughput, but 3GPP standardised in Release 13 two specific variants to specifically address new IoT use cases, allowing large-scale IoT deployments and a reduction in the device complexity: LTE-M, also known as enhanced MTC (eMTC), and NB-IoT. While the former technology was defined to operate in regular LTE deployments, using the smallest possible channel size (i.e. 1.4 MHz), the latter was designed to operate in a very small 180 kHz channel size, which allowed it to be deployed in standalone mode (typically reusing GSM channels), in regular LTE bands, or within LTE guard bands. Generally, LTE-M is richer in capability than NB-IoT; as defined in its initial release, LTE-M can support mobility, voice over LTE (VoLTE), and a data rate up to 1 Mbps, while NB-IoT is limited to 30 kbps. On the other hand, NB-IoT achieves theoretically better coverage and lower power consumption. Both technologies are now deployed extensively all over the world while they keep evolving and improving in subsequent 3GPP releases. The evolution of LTE-based IoT features as well as primary key KPIs and design objectives of the respective Cat-M and Cat-NB devices are discussed in detail within Section 2.2.1.

At the dawn of 5G, The International Mobile Telecommunications in the context of 5G (IMT-2020), issued by the International Telecommunication Union (ITU), defined the minimum requirements [13] and evaluation guidelines [14] to identify the candidate IMT-2020 radio interface technologies when addressing three 5G usage scenarios: enhanced Mobile Broadband (eMBB) communications, Ultra-Reliable and Low Latency communication (URLLC) and massive machine-type communication (mMTC). Given those performance requirements, 3GPP considered and verified both LTE and NR as potential cellular technologies able to fulfil the objectives of IMT-2020, meaning that they are compliant with the level of performance established by the ITU for IMT-2020.

Table 1. Minimum technical requirements for IMT-2020. Source: [14]

Technical requirements	eMBB	URLLC	mMTC
Peak data rate	DL: 20 Gbps UL: 10 Gbps		
Peak spectral efficiency	DL: 30 bit/s/Hz UL: 15 bit/s/Hz		
User experienced data rate	DL: 100 Mbps UL: 50 Mbps		
Area traffic capacity	10 Mbit/s/m^2		
User plane latency	4 ms	1 ms	
Control plane latency	20 ms (10 ms encouraged)	20 ms (10 ms encouraged)	
Connection density			1.000.000 devices/km^2
Reliability		$1\text{-}10^{-5}$ success	

Co-funded by the Horizon 2020 Framework Programme of the European Union

		probability	
Mobility	Up to 500 km/h (for high-speed trains)		
Mobility interruption time	0 ms	0 ms	
Bandwidth*	Minimum 100 MHz and up to 1 GHz for higher frequency bands (e.g. millimetre wave)		
(*) General/Non-specific requirement.			

Figure 2.3 shows the evolution of C-IoT air interfaces throughout the 3GPP releases, where "SI" means Study Item and "enh." refers to enhancements. Even if Release 15 paved the way to the new 5G-oriented NR air interface, 3GPP also addressed in parallel new enhancements for LTE. Thus, the label "5G" actually refers to the study and work performed in Release 15 and beyond, including the scope of both NR and LTE.

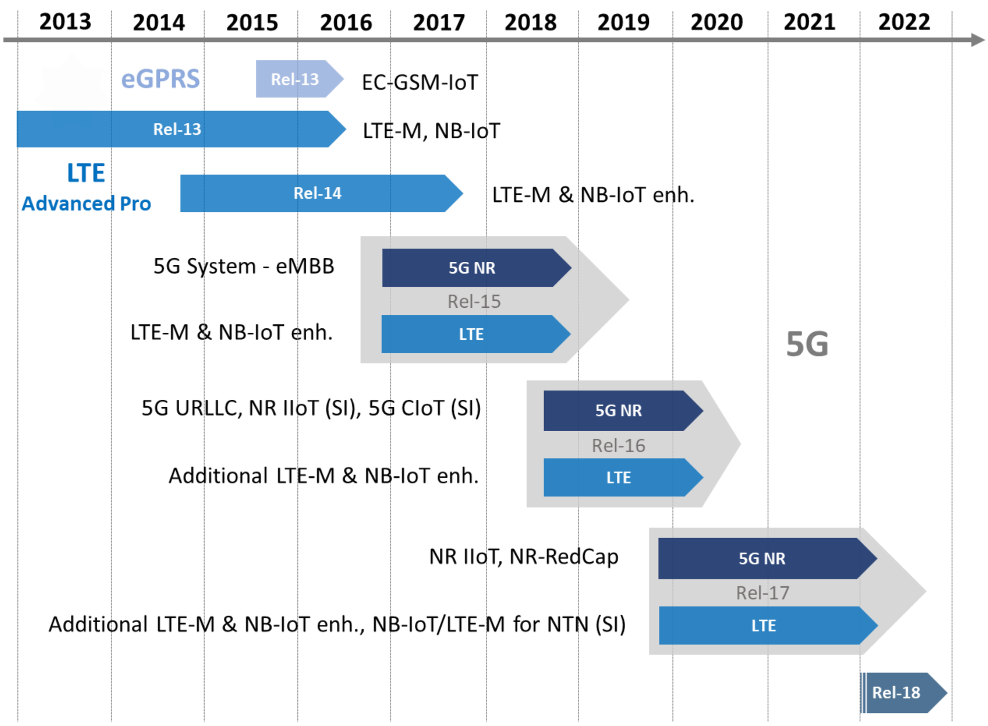


Figure 2.3. C-IoT evolution throughout 3GPP releases

While the primarily focused 5G NR usage scenario, eMBB, was defined by the 3GPP in the first phase of the 5G System (5GS) in Release 15, URLLC was mainly defined within the Release 16. Thanks to the capabilities of both usage domains, the NR air interface is capable of addressing new IoT applications aimed at efficiently supporting advanced M2M communication. To go a step further in the context of IoT, NR leverages the capabilities of eMBB and URLLC focused features to enable the ubiquitous connectivity in industrial applications and verticals, becoming the boost the next wave of industrial transformation needs. In that direction, the study on NR Industrial Internet of Things (NR IIoT) started in Release 16, but the standardization work will be further completed in Release 17. This new communication class aims to cover industrial applications related to factory automation (i.e. logistics, sensor networks, robotics, and augmented reality) where both

Co-funded by the Horizon 2020
Framework Programme of the European Union

eMBB and URLLC features become vital elements to support high transmission reliability and performance. In addition, TSN is a key enabler for NR IIoT. It encompasses a set of standards identified by the IEEE 802 family of standards that enables Ethernet wired networks to ensure Quality of Service (QoS) features for time-sensitive traffic and critical-data applications, in order to provide deterministic transmissions by synchronizing various equipment components to a single master clock [15]. In TSN, it is not necessary to use the internet protocol since Ethernet frames can be transported over the 5G system in an Ethernet power distribution unit (PDU) session type. Mechanisms to ensure deterministic delays and synchronization were defined by IEEE and the objective of 3GPP was to adapt these mechanisms to the wireless and 5G world.

On the other hand, 5G mMTC is the answer to serve the mainstream IoT applications. This usage domain is tailored to allow a high-density connection of low-complexity power-limited devices such as sensors and asset trackers. In this scenario, neither the demand of high data rates nor the latency budget is stringent, so throughput and latency step aside to long-lasting battery devices with wireless connectivity at a lower cost. Given the IMT-2020 minimum requirements setting the connection density as the main objective for the mMTC usage scenario (see Table 1), 3GPP proposed LTE-based IoT solutions (LTE-M and NB-IoT, Release-15 and above) to address mMTC, which was accepted by the ITU.

Finally, the currently ongoing work of 3GPP in Release 17 related to C-IoT is not limited to provide enhanced IIoT support for NR and further improvements of LTE-M and NB-IoT, but it is also committed to introducing new communication paradigms. This is the case of the current studies on LTE-M/NB-IoT over NTN and NR-RedCap. The progress and evolution of NR-based IoT is discussed in detail within Section 2.2.2.

All in all, the concept of 5G IoT encompasses the evolution towards novel IoT applications and use cases. The capabilities of eMBB and URLLC allow verticals to explore beyond todays' cellular protocols and open doors to new radio access solutions like NR-RedCap, NR IIoT, and IoT over NTN, especially suited for upcoming IoT scenarios. In any case, the objective of NR-based IoT is not to overlap the use cases already covered by LTE-M and NB-IoT, but to expand their scope and address the uncovered IoT-related scenarios.

2.2.1 LTE-based IoT

As seen in previous section, 3GPP has some well-established technology solutions to address IoT market, especially towards the low-power wide-area network services (which more or less overlap with the massive IoT category). After their initial specification in Release 13 (early 2016), LTE-M and NB-IoT were improved in subsequent Releases, 14 and 15. As a result, when 3GPP submitted its proposal to ITU for 5G (to address eMBB, URLLC, and mMTC, the so called 5G triangle of services), it submitted NR for eMBB and URLLC, while LTE-M and NB-IoT were accepted as already meeting the requirements for the mMTC aspects of IMT-2020 [16].

Nevertheless, further RAN enhancements were deemed necessary for supporting larger number of services and use cases as well as for coexistence

Co-funded by the Horizon 2020 Framework Programme of the European Union

with NR. In addition, since a new 5GS was specified in Rel-15 [17] including 5G RAN, also called Next-Generation RAN (NG-RAN, i.e. the evolved RAN architecture with the 5th generation of mobile networks), the 5G core (5GC) network needed to be developed to support 5G mMTC requirements. To this end, 3GPP undertook work in the following topics within Rel-16: a) *Additional LTE enhancements for MTC* [18] and *NB-IoT* [19] to support their ongoing commercial and deployment growth by further improving their respective network operation and efficiency in a range of areas including coexistence with NR and connection to 5GC; b) *Cellular IoT support and evolution for the 5G system study* [20] on the efficient support of C-IoT capable UEs in 5G core network and architectural alternatives was concluded in TR 23.724 and respective work item [21] was tasked to specify 5GS enhancements and enable C-IoT functionalities for 5GS capable devices that support eMTC and/or NB-IoT.

Even within the ongoing Rel-17 specification, 3GPP has kept allocating work to LTE-based C-IoT enhancements. *Rel-17 enhancements for NB-IoT and LTE-M* work item [22] aims to bring enhanced or new features on the existing LTE technologies for IoT to support their long-term lifecycle and address lessons drawn from deployments and trials, but also to broaden use cases for legacy cellular IoT.

Several works exist in literature describing the LTE-based IoT technologies and their enhancements over the past 3GPP Releases, for example see [5][23][24] as well as [25][26][27] directly from 3GPP. In the following, we first present the respective devices, KPIs, and design objectives of LTE-based IoT, and then summarise the key aspects and features introduced through 3GPP releases.

2.2.1.1 Devices, KPIs and design objectives

The devices supporting LTE-M and NB-IoT technologies are denoted, respectively, as Cat-M and Cat-NB. Each of the two device categories has two variants, Cat-M1/M2 and Cat-NB1/NB2. Furthermore, all devices are fully backward compatible with Release 13, meaning that the new features introduced in subsequent releases for LTE-based IoT can be introduced gradually.

Generally, Cat-M and Cat-NB can be seen as stripped-down versions of regular LTE devices, with the design target being low cost, improved (indoor) coverage, and very long battery life as needed for battery-powered IoT applications, including utility meters, wearables, alarm panels, and asset trackers. Primary design objectives of Cat-M and Cat-NB devices have been:

- **Reduced cost, smaller footprint**. While regular LTE uses two antennas on the device side, only one antenna is possible in LTE-M and NB-IoT designs, simplifying signal processing. The use of smaller channel sizes further simplifies processing. Another cost reduction is enabled by the elimination of the duplexer (the specific filter that protects the receive path from the transmit signal) which is not needed in half-duplex frequency division duplex (HD-FDD), the mode which allows the UE to transmit and receive data on different frequencies, but not at the same time, and which is used mostly in

Co-funded by the Horizon 2020 Framework Programme of the European Union

LTE-M and NB-IoT. This allows a drastic simplification of the radio front-end, allowing a single hardware design to operate globally.

- **Improved coverage**. Removing one antenna negatively impacts receiver sensitivity, so in order to compensate for this loss and to improve the coverage (as necessary for deep indoor deployments such as smart meters), coverage enhancement (CE) modes were introduced. CE modes (i.e. CE mode A optimized for moderate coverage enhancement and CE mode B providing extremely deep coverage) are simply signal repetitions, and although not an optimal solution from an information theory standpoint, they are a low-cost technique for improving signal-to-interference-plus-noise ratio (SINR).

- **Very long battery life**. New power saving schemes, namely, extended Discontinuous Reception (eDRX) and Power Saving Mode (PSM), and protocol optimizations were introduced allowing IoT devices to enter into a deep sleep mode as fast as possible and as long as possible, resulting in a reduction of power consumption to the lowest level possible.

Table 2 presents the main KPIs that define the EC-GSM-IoT, LTE-M and NB-IoT technologies, all of them targeting applications with dense connectivity requirements.

We should also note that, in addition to Cat-M/NB, regular LTE device variables, I.e. *Cat-1*, *Cat-4*, etc., can be considered to deliver the higher-end broadband IoT. These devices are definitely of higher cost compared to Cat-M and Cat-NB devices but can deliver multi-Mbps data rates. An interesting LTE-based variant is *Cat-1bis* device which is essentially a Cat-1 capability device but with only 1 Rx chain implementation to address market demand for low cost, low power consumption, and smaller form factor IoT solutions, for e.g. wearables that may but may typically require relatively high data rates and flexibility in terms of baseband implementation complexity. Even if 2 Rx chains are used in such devices, the actual performance will be close to 1 Rx in several cases, since the separation between the antennas will be very small. Very small form factor solutions with 1 Rx can be used. Back in 2014, 3GPP defined requirements for MTC UEs (Cat-0) with 1 Rx chain [46]. But Cat-0 UEs were never deployed as in the end no meaningful outcomes were defined for the wearables market; the main shortcoming was that very low data rates were supported while legacy-like supported bandwidth (i.e. 20 MHz) did not allow for significant modem complexity reduction anyway. On the other hand, Cat-1 UE fulfilled these data rate (and also mobility) requirements. To this end, in 2017, 3GPP developed requirements for Cat-1 UE with one receive chain [47]. Cat-1bis can still connect to legacy networks but it can also get the benefits of key enhancements and features (e.g. Rel-14 VoLTE support) in networks that recognize this UE (can start from Rel-13).

Table 2. EC-GSM-IoT, LTE-M and NB-IoT KPIs

KPIs	**EC-GSM-IoT***	**LTE-M**	**NB-IoT**
Spectrum	In-band GSM	In-band LTE (1.4 MHz)	In-band LTE Guard-band LTE

Co-funded by the Horizon 2020 Framework Programme of the European Union

			Standalone (200 kHz)
Peak Data rate	350 bps - 70 kbps (Gaussian minimum-shift keying - GMSK) Up to 240 kbps (Eight Phase Shift Keying - 8PSK)	**Cat-M1** Up to 1 Mbps (DL & UL, Full Duplex FDD) **Cat-M2** 4 Mbps (DL) 7 Mbps (UL)	**Cat NB1** 26 kbps (DL) 66 kbps (UL, multi-tone) 16.9 kbps (UL, single tone) **Cat NB2** 127 kbps (DL) 159 kbps (UL)
Bandwidth	200 kHz	1.08 MHz (Cat-M1) 5 MHz (Cat-M2)	200 kHz
Battery life	Long (<10 years)	Long (<10 years)	Very long (10 – 15 years)
Latency	700 ms – 2 s	10 - 15 ms	1.5 – 10 s
Complexity/ Cost	Low as compared to GPRS/GSM devices	Low as compared to LTE Cat-1/Cat-4 devices	Ultra-low as compared to LTE Cat-1/Cat-4 devices
Coverage	154 dB MCL with PC3*** 164 dB MCL with PC5***	>155.7 dB MCL**	164 dB MCL for standalone

(*) Not deployed.

(**) MCL= *Maximum Coupling Loss*, defined as the maximal total channel loss between user device and base station antenna ports at which the data service can still be delivered.

(***) PC = *Power Class*, UE classification based on the maximum transmit power (23 dBm for PC3, 20 dBm for PC5)

2.2.1.2 Release 14

The next release of the two LTE-based technologies for massive IoT, though coming a year after their initial release, improved their specified basic functionality in several aspects:

Data rate enhancements

For LTE-M, the data rate improvement came from the introduction of several features. Primarily, by supporting a new device category, Cat-M2, with increased bandwidth for data channels (5 MHz instead of 1.4 MHz) and increased maximum Transport Block Size (TBS) in DL and UL. Option of maximum UL TBS was also supported for Cat-M1, to increase its UL peak data rate. Furthermore, assistance signaling is introduced to allow devices indicate their preferable maximum bandwidth configuration in a CE mode. Support of up to 10 DL hybrid automatic repeat request (HARQ) processes in FDD (instead of 8) and of HARQ-acknowledgement (HARQ-ACK) feedback

Co-funded by the Horizon 2020 Framework Programme of the European Union

bundling in HD-FDD (to transmit HARQ-ACK feedback transmission for multiple DL data transmissions in a subframe) gave a boost to DL peak data rates. Finally, it became possible for the device to indicate capability of faster frequency retuning, resulting into smaller guard period. For NB-IoT, a new device category, Cat-NB2, was also introduced with increased TBS size and code rate, but also with support for two HARQ processes in order to address the existing issue of high delay due to UL and DL scheduling gaps.

Voice, Coverage and Mobility enhancements

Several features were also introduced to enhance experience and coverage of real-time, delay-sensitive applications such as VoLTE for legacy and Cat-M devices, especially in HD-FDD, since previous Release was optimised only for delay-tolerant cases. First, new physical uplink shared channel (PUSCH) repetition factors - and respective downlink control information (DCI) fields' update - were introduced to match the VoLTE traffic pattern (i.e. 20 ms periodical speech frame). In addition, possibility was supported to restrict UL and DL data transmissions to QPSK modulation to improve link performance in case of repetitions. Dynamic control (from network-side) of the HARQ feedback delay (instead of fixed 4 ms) was also introduced via DCI field update to work together with the aforementioned HARQ-ACK bundling feature and allow more efficient scheduling of VoLTE transmissions. In addition, to improve UL and DL link adaptation, support of CE through repetition was also introduced for sounding reference signals via the use of the UL part of the special subframe in time division duplex (TDD) mode. Finally, again for LTE-M, mobility support in connected mode was enhanced by introducing intra-frequency reference signal received quality (RSRQ) - in addition to existing reference signal received power (RSRP) – measurements, as well as inter-frequency RSRP/RSRQ measurements to better support the mobile wearables and VoLTE use cases.

Capacity enhancements

To support the 5G requirement of a million devices per Km^2 in NB-IoT, Release 14 introduced random access on non-anchor carriers (i.e. the secondary carriers in multicarrier operation of NB-IoT, which do not carry physical channels for initial cell selection) for improved total number of system random access opportunities, which theoretically increased the system capacity by more than a 10-fold. In addition, paging to non-anchor carriers was supported for mobile terminated reachability for a very high number of users.

Power enhancements

A new, 14 dBm, power class is introduced (in addition to existing 20 and 23 dBm) for NB-IoT. The motivation from this introduction was to reduce the power amplifier drain current and enable the use of simpler and more compact battery types for NB-IoT devices.

Positioning

For LTE-M, performance requirements were introduced for the already existing Enhanced Cell ID (E-CID) positioning method, to be based not only cell identity, but also on estimated TA that determines the round-trip time,

Co-funded by the Horizon 2020 Framework Programme of the European Union

hence distance, between device and evolved Node B (eNB, the base station equivalent in LTE networks). Also enhancements were introduced for Observed Time Difference of Arrival (OTDOA) method, to be based on device measuring the time of arrival (ToA) on a set of DL positioning reference signals. In addition, OTDOA enhancements were supported in the sense of positioning reference signal (PRS) configurations to match the positioning accuracy of legacy LTE; mainly regarding time/frequency mapping, to adjust to the limited bandwidth / longer duration nature of the technology, but also for frequency hopping support to obtain frequency diversity gains. For NB-IoT, support of positioning through E-CID and OTDOA was also introduced considering respective new channels and signals, e.g., the NB-IoT positioning reference signal (NPRS).

Multicast transmission

For both LTE-based IoT technologies, support for multicast transmission was introduced to, e.g., provide group software updates. This so called Single Cell Point-to-Multipoint (SC-PTM) transmission was introduced based on the existing Multimedia Broadcast Multicast Service framework and only supported in idle mode. A new system information block (SIB) was introduced which can contain scheduling information for one Single Cell Multicast Control Channel (SC-MCCH) per cell, and SC-MCCH can contain scheduling information for one Single Cell Multicast Traffic Channel (SC-MTCH) per multicast service.

2.2.1.3 Release 15

Release 15 enhancements for LTE-based IoT technologies were completed in 2018 with key focus on supporting new use cases (such as high speed devices, small factor and lower power consumption wearables), improving latency, spectral efficiency, power consumption as well as building on feedback from early deployments.

Latency reduction

To improve latency for LTE-M, the focus was primarily on reducing system acquisition time. This was achieved by: i) improving initial cell search via E-UTRA absolute radio frequency channel number (EARFCN) pre-provisioning, ii) introducing a new, more densely transmitted reference signal than the primary and secondary synchronization signals, the so called Resynchronization Signal (RSS), for device to re-acquire time and frequency synchronization faster (improving also of course the device power consumption), iii) improved master and system information block (MIB and SIB) acquisition performance, and iv) a new indication of whether MIB (and SIB, respectively) information has been updated recently to avoid frequent re-acquisition. Also, support of Early Data Transmission (EDT, i.e. data transmission occurring within the random access procedure) was introduced to help with latency reduction. For NB-IoT, support for radio link control (RLC) unacknowledged mode (UM) is added (complementing the existing acknowledged mode and transparent mode) to reduce the need to send RLC signalling over the air for delay tolerant/recoverable IoT traffic.

Co-funded by the Horizon 2020 Framework Programme of the European Union

For NB-IoT, system acquisition time is reduced via support of 16 (instead of 8) SIB1-NB – the primary NB-IoT system information block – repetitions on anchor and non-anchor carriers for faster SIB1-NB decoding.

Capacity enhancements

Higher-order modulation in DL data unicast transmissions, i.e. 64 quadrature amplitude modulation (QAM), and associated channel quality indication (CQI) table with larger range is supported to increase DL spectral efficiency. For the UL case, finer-granularity UL data channel resource allocation is introduced, i.e., sub- physical resource block (PRB); 3 or 6 subcarriers. Furthermore, DL/UL data resource allocation with a more flexible starting PRB (instead of restricted by 6-PRB narrowbands) is introduced for devices configured in CE mode with max 1.4 MHz. Finally, support of frequency domain cell-specific reference signal (CRS) muting can be indicated by devices to network for reducing inter-cell interference.

Power enhancements

Apart from savings in synchronization phase via RSS, reduced device power consumption is also achieved in this Release via reduced device downlink monitoring, reduced signalling and reduced uplink transmission. More specifically, introduced features for LTE-M/NB-IoT include: i) EDT for data transmit in the third message transmission (Msg3) and possible data receive already in fourth message transmission (Msg4) of random access procedure, for cases where IoT device needs to transmit small amount of data (~100s bytes), ii) Wake-up signal (WUS), a compact signal transmitted before (up to 2 seconds) the paging occasion of a UE supposed to be in idle mode (DRX or eDRX), allows the device to skip paging procedures and to go to a near sleep, very low-power state If WUS is not detected, iii) relaxed neighbouring cell measurements (up to 24 hours) for cell reselection which can be really useful to stationary UEs which suffer from bad coverage but not inter-cell interference. Specifically for LTE-M, HARQ-ACK feedback for UL data via DCI is supported to enable early termination of DL monitoring or of uplink transmission at the device. Furthermore, in case of the newly introduce sub-PRB UL allocation with 3 subcarriers, a new π/2- binary phase shift keying (BPSK) modulation, using one subcarrier at a time, can be used to achieve near 0 dB baseband peak-to-average power ratio (PAPR), which can be beneficial for device power consumption (as well as UL data coverage). For NB-IoT, the granularity of Power Headroom Report (PHR) transmitted in Msg3 is also increased to have 16 levels. Device is also allowed to quickly release Radio Resource Control (RRC) connection, by considering successfully acknowledged the receipt of the *RRC Connection Release* message, as soon as the UE has sent HARQ-ACK (instead of waiting up to 10 seconds). In addition, scheduling request procedure is improved, allowing Buffer Status Report (BSR) reporting configuration and resource activation/release by network for higher power (and resource) efficiency.

Access control

For both technologies, a new access barring mechanism is introduced, to enable network bar device access per coverage-level support. This approach is to help with high-load situations by temporarily barring devices with e.g. highest CE levels, thus, more repetitions and highest resource consumption.

Co-funded by the Horizon 2020 Framework Programme of the European Union

NB-IoT enhancement

Several aspects are introduced to NB-IoT to bring it up to speed with legacy LTE releases, improve its operation and allow new deployment options. Support for TDD is introduced with all LTE configurations for UL/DL (except for 0 and 6) and special subframe supported. In addition, eNB lower power classes are defined for NB-IoT to allow small cell deployment while, on the other hand, new NB-IoT physical random access channel (NPRACH) format is introduced - 1.25 kHz subcarrier spacing (SCS), 800 µs cyclic prefix (CP) - together with frequency hopping, to allow unambiguous range determination in cells of up to 120 km. UE differentiation is made possible by enabling network to collect and store information about the UE and its traffic profile. Last, but not least, standalone anchor/non-anchor carriers are allowed to be configured with in-band and guard-band non-anchor/anchor carriers; this feature allows for small slices of non-LTE spectrum to be used as a standalone NB-IoT carrier and be linked with NB-IoT carriers associated to LTE spectrum.

2.2.1.4 Release 16

This release, arrived in 2020, built on feature introduced in earlier releases, mainly for further improving network operation but also for improved efficiency in several aspects.

Power and Transmission efficiency enhancements

WUS feature was enhanced to group WUS (GWUS) allowing wake up a configurable group of UEs (based on UEs paging probability and/or ID) rather than all UEs monitoring the same paging opportunity (PO). Mobile terminated EDT access (MT-EDT, in addition to previously introduced mobile originated version) is supported to allow data transmission within the random access procedure triggered in response to a paging message. Also, transmission of UL data has been further enhanced by using eNB preconfigured uplink resources (PUR) to omit both the random-access preamble transmission (Msg1) and the random-access response (Msg2), given the TA has been evaluated as valid by UE. Moreover, possibility to schedule multiple transport blocks (TBs), unicast and multicast, with a single DCI is introduced; this is to reduce PRBs overhead spent for control transmission, as well as guard time subframes in HD-FDD, and improve resource utilization.

In addition, for LTE-M, a new type of DL quality reporting is introduced representing the required number of repetitions at the MTC physical downlink control channel (MPDCCH) for reliable reception, thus, reflecting MPDCCH quality instead of physical downlink shared channel (PDSCH) quality. Also, to improve control channel decoding performance, it is specified how to use not only demodulation reference signal (DMRS) but also CRS for MPDCCH demodulation. For NB-IoT, eNB can configure an idle UE for DL channel quality report in Msg3 for non-anchor access to avoid mismatch with coverage level in cells with interference. Further, eNB is allowed to transmit NRS to page UE on non-anchor carrier, even without paging NB-IoT physical downlink control channel (NPDCCH) transmitted,

Co-funded by the Horizon 2020 Framework Programme of the European Union

and UE can perform serving cell measurements on such non-anchor paging carrier.

Mobility, CE mode and latency enhancements

For LTE-M, signaling of RSS configurations for neighbor cells is introduced to improve intra-frequency RSRP measurements performance for neighbour cells in both idle and connected mode. Furthermore, a Radio Resource Management (RRM) measurement relaxation is introduced allowing UEs to meet the requirements using a longer, configurable, measurement cycle to fully utilize possible power saving gain from WUS (instead of frequent wake up for measurements). For NB-IoT, assistance information is supported from/to E-UTRAN and to GERAN for idle mode inter-RAT cell selection. In addition, support for UE-specific DRX is introduced, where UE monitors paging according to the max between the cell default DRX value and an eNB-broadcasted minimum UE-specific DRX value, to reduce paging latency.

Standalone LTE-M deployment enhancement

On carriers that are not used for normal LTE, it is now possible to transmit MPDCCH/PDSCH to UEs in CE modes A or B in the “LTE control channel region", i.e. the first 1-3 orthogonal frequency division multiplexing (OFDM) symbols in each DL subframe previously assumed to be occupied by LTE control channels for normal LTE UEs. This feature increases the transmission efficiency from 7% to 14% depending on carrier bandwidth.

NR coexistence enhancement

Rel-15 NR was designed to already support spectrum sharing with legacy (Rel-13/14/15) LTE-based IoT; the radio frequency (RF) coexistence aspects have been described in [28]. Release 16 introduces some enhancement features to further improve networks performance when they coexist.

For LTE-M, finer-granularity resource reservation, with configurable patterns, is introduced in time (subframe, slot, or symbol level granularity) and frequency (LTE resource block group level granularity) for unicast control and data transmissions in connected mode in CE mode A/B. Furthermore, 1 or 2 DL subcarriers’ puncturing is supported (at the lower or higher edge of each 6-PRB narrowband) to achieve PRB alignment between LTE-M and NR, affecting slightly the MPDCCH/PDSCH transmissions in connected mode in CE mode A/B.

For NB-IoT, coexistence (in NR in-band, NR guard-band, or standalone operation) is enhanced by allowing configuration of the DL/UL resource reservation in subframe/slot/symbol-levels on non-anchor carriers for unicast transmission to avoid resource overlapping with NR channels and signals. Also dynamic indication is allowed on whether the resource reservation is applied or not.

Connection to 5G core (5GC)

In addition to aforementioned features, an important advancement of this release is the introduced support for connection of Cat-M and Cat-NB UEs to 5GC network, using LTE as baseline. RRC inactive mode is supported and

Co-funded by the Horizon 2020 Framework Programme of the European Union

additionally the User Plane (UP) C-IoT 5GS optimisation is supported in RRC idle mode. Some LTE-based IoT features (EDT, PUR) are supported only in idle mode using the UP-optimisation solution and are not supported in inactive mode. On the other hand, long extended DRX in idle mode is supported, and RAN paging cycles of 5.12 and 10.24 seconds are supported in inactive mode.

2.2.1.5 Release 17

Release 17 enhancements for LTE-M and NB-IoT are ongoing and expected to be completed by December 2021. For NB-IoT, enhancements to be specified include increased data rate via 16 QAM, support for carrier selection and reduced time for RRC re-establishment. For LTE-M, planned enhancements include increase on HD-FDD DL spectral efficiency via 14 DL HARQ processes and possible transmissions power reduction for UEs supporting PUSCH sub-PRB resource allocation.

In parallel, a Rel-17 *NB-IoT/eMTC over NTN* study targets to provide recommendations for enhancing NB-IoT and LTE-M to address IoT operation in remote areas utilizing satellite connectivity. The study will be based on prior NR NTN study and conclusions in TR 38.821 [66] and on the also ongoing Rel-17 work item on solutions for NR NTN. The objectives include the identification of scenarios applicable to NB-IoT/eMTC as well as the study of enhancements for these scenarios. Enhancements are expected in aspects such as random access procedure/signals, mechanisms for time/frequency adjustment, tracking area, etc. More information regarding the aforementioned NTN-related works within 3GPP can be found in Section 2.2.2.3.

2.2.2 NR-based IoT

As we have seen at the start of the chapter, 5G NR opens the door to allow for communication of much more sophisticated and higher data rate objects that must also meet stricter requirements in terms of latency and reliability. These more demanding objects are sometimes referred to as industrial IoT or critical IoT objects in order to distinguish them from low profile IoT objects, called massive IoT (generally mapping to the mMTC requirements). Section 2.2.2.1 presents the existing 3GPP work on NR IIoT, the targeted use cases and KPIs, as well as the aspects considered in relation to TSN for enabling NR IIoT.

Furthermore, 3GPP current efforts within Release 17 for emerging C-IoT solutions include the study and specification of RedCap NR devices. The ongoing NR-RedCap work is targeting to enable devices of much lower cost (~40-50%) from legacy (Rel-15/16) NR devices that have been developed to address eMBB and/or URLLC services. The upcoming RedCap devices can be viewed as NR-based C-IoT devices targeting to address three specific use cases (see Section 2.2.2.2) with requirements that may not be adequately addressed so far by the aforementioned LTE-based C-IoT solutions.

Figure 2.4 illustrates the scope of NR-RedCap and NR IIoT in the framework of the mainstream 5G usage scenarios (i.e. eMBB, URLLC and mMTC). Whilst NR IIoT benefits from the capabilities of eMBB and URLLC, the specifications

Co-funded by the Horizon 2020 Framework Programme of the European Union

of RedCap NR devices are somehow in between of eMBB and mMTC, particularly, ranging from LTE Cat-1 to Cat-4. The figure includes the main features and application examples of each radio access technology (RAT).

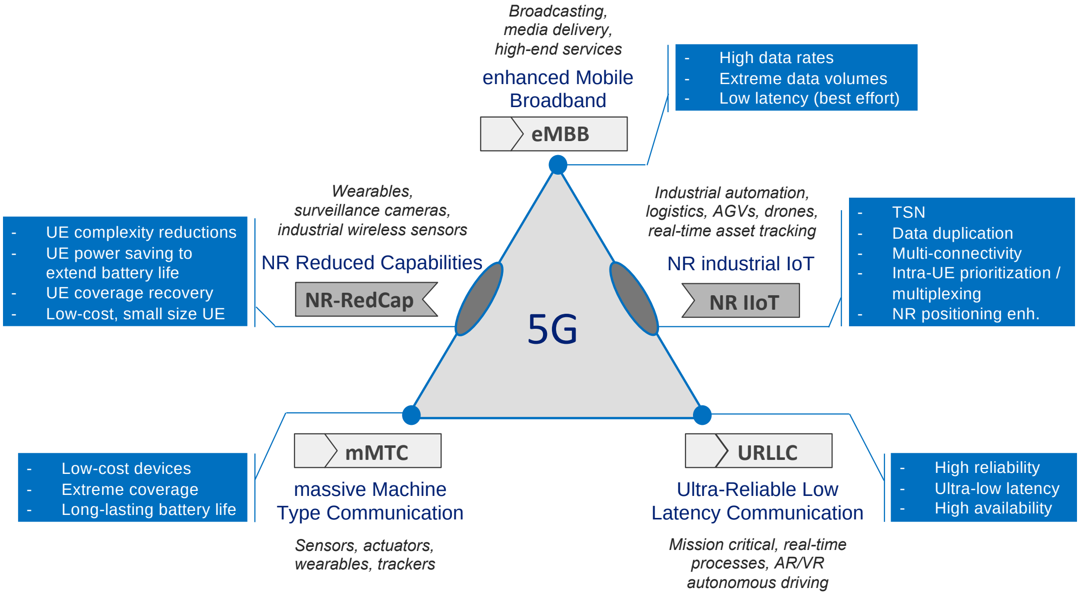


Figure 2.4. 5G usage scenarios

In parallel, 3GPP 5G-era efforts have been also encompassing the study and specification of satellite-based communications in order to address the growing demand for new emerging services, such as maritime or railway, that cannot be handled by current terrestrial systems capabilities. Adaption of NR features to support NTN is being specified while the feasibility of ensuring NTN connectivity of LTE-IoT devices is being studied. This NTN work is also of particular interest to iNGENIOUS considering that a subset of use cases of interest is expected to rely on satellite-based connectivity. To this end, Section 2.2.2.3 presents the existing 3GPP considerations on NTN targeted use cases, deployment options, architecture and channel modelling, as well as the various air interface design technical aspects defined and ongoing planned work.

2.2.2.1 NR Industrial IoT

Although originally MTC was considered only for low data rates devices and applications, 5G NR brings significant improvements in latency and data rate compared to 4G, and these improvements are key in meeting the strict requirements in vertical markets such as factory automation (industry 4.0), transport, energy, or entertainment, including augmented and virtual reality. Most of these improvements are defined within the context of the URLLC side of the 5G triangle. URLLC services are enabled by the flexible frame structure (allowing a very short transmission time interval), pre-emptive scheduling, and anticipated re-transmission for fast turn-around, grant-free transmission, etc.

To address these services and use cases of strict requirements 3GPP worked on the topic of enhanced URLLC (*eURLLC*) [29] which was tasked to specify Physical Layer Enhancements for NR URLLC - including enhancements in

Co-funded by the Horizon 2020 Framework Programme of the European Union

PDCCH, uplink control information (UCI), PUSCH, scheduling/HARQ, inter UE Tx prioritization/multiplexing and UL configured grant transmission - for broadening the spectrum of supported use cases with tight latency and reliability requirements.

In parallel, embracing the potential of NR eMBB and URLLC as a launch pad for industrial automation became a major concern in Release 16. Therefore, the 3GPP started a Study Item (SI) on NR enhancements to URLLC and IIoT [30] followed by *NR IIoT* work item [31]. NR IIoT introduced improved reliability thanks to enhanced packet data convergence protocol (PDCP), an upper layer of the protocol stack, via duplication, mechanisms to prioritize traffic between UEs and within a UE, enhancements required to support identification and selection of non-public networks, and a means to support time sensitive networking TSN.

This NR-based type of IoT is further strengthened in Release 17 and it is expected to fully support the most stringent requirements of critical and industrial connected objects. More specifically, the currently ongoing *Enhanced IIoT and URLLC support* work item [32] has a key objective to investigate required enhancements for supporting unlicensed operation, especially in controlled environments with only devices operating on the unlicensed band are installed by the facility owner and where unexpected interference from other systems and/or radio access technology only sporadically happens.

NR IIoT targets the following use cases: (i) industrial automation environments requesting higher level of flexibility, productivity and lower installation and maintenance cost, (ii) intelligent manufacturing and supplies management, (iii) streamlining logistics by using, for instance, AGVs, drones, sensor networks or self-driving forklifts, and (iv) real-time end-to-end asset tracking (e.g. trucks, containers, goods) across the supply chain either to know their exact position avoiding potential losses or, even, to optimise routes and generate predictive data patterns.

Industrial environments have particular demands in terms of: (i) high reliability to react to critical events and to allow the best positioning accuracy of the devices, (ii) low latency to respond to real-time decisions, and (iii) guaranteed bandwidth to cover IIoT specific use cases and to ensure the connectivity across the industrial premises.

For this purpose, the role of TSN, as well as data duplication and multi-connectivity enhancements and intra-UE prioritization/multiplexing, are fundamental. In TSN use cases, the UEs need to manage a mixture of the following different traffic: (i) multiple periodic streams of different periodicities coming from different applications, (ii) aperiodic priority traffic as a result of critical events such as alarms or safety detectors; and (iii) best effort type of traffic such as eMBB traffic.

Three mainstream requirements are associated to TSN: (i) reliability, (ii) latency and (iii) time synchronization accuracy. On one hand, the target reliability is 99,99% and 99,9999% since PDCP data duplication (i.e. higher layer redundancy) is considered to increase the reliability. Concerning the latency, 0.5 ms one-way latency target for both DL and UL is considered (grant-free scheduling for UL, and both TDD and FDD modes). Finally, time

Co-funded by the Horizon 2020 Framework Programme of the European Union

synchronization accuracy over the Uu interface (i.e. air interface between the gNB - the base station equivalent in next generation networks - and the UE) depends on the maximum gNB-to-UE distance in case the UE would not compensate for the radio propagation delay of the channel. The achievable time synchronization accuracy between the gNB and TSN Grand Master clock can be much less than 1µs.

On the other hand, the 3GPP has identified in [30] PDCP duplication and higher layer multi-connectivity aspects, such as assessment of gains of duplication with more than two copies, potential enhancements to achieve resource efficient PDCP duplication, and captures RAN aspects of higher layer multi-connectivity solutions. Also, the intra-UE traffic prioritization and multiplexing, data and control channels, different latency and reliability requirements and different types of resource allocations for both uplink and downlink directions are considered.

From the architecture perspective, Non-Public Networks (NPNs), also known as private networks, allow to deploy the 5G system for private use. NPNs are exclusive mobile networks that let operators manage the resources of their own network so that the connected elements are part of a closed network infrastructure. This is highly beneficial not only for the industry sector, but also in many other scenarios. Resources are not shared with other users, hence guaranteeing the autonomy of verticals in the 5G framework. Another important consideration in industrial IoT scenarios is the network security. In this regard, 5G NPNs are a powerful tool to provide enterprises with a superior level of security as compared to the one provided over a public network. NPNs can be deployed in two different ways: as a standalone NPN (SNPN), or as a public network integrated NPN (PNI-NPN) where the NPN relies on the network functions provided by a PLMN.

5G Local Area Networks (LAN) in industrial environments are also a significant improvement from the infrastructure point of view. 5G LANs aim to provide services with similar functionalities as LANs, but enhanced with 5G capabilities (e.g. high performance, long distance access, mobility, customization features and security). As defined in Release 16, the 5GS is evolved to offer private communication using Internet Protocol (IP) and/or non-IP for UEs that are members of a 5G virtual network group [33]. The integration of LAN features and the 5G Core network is considered a cornerstone of Private Networks.

Finally, 3GPP Release 16 has defined a set of network components, i.e. Network TSN Translator (NW-TT) and Device Side TSN Translator (DS-TT) in order to integrate mobile devices connected to 5GS to become native TSN devices [34]. Thus, specific functions to integrate mobile devices as seamlessly part of fixed TSN infrastructure are in the NW-TT and DS-TT.

Co-funded by the Horizon 2020 Framework Programme of the European Union

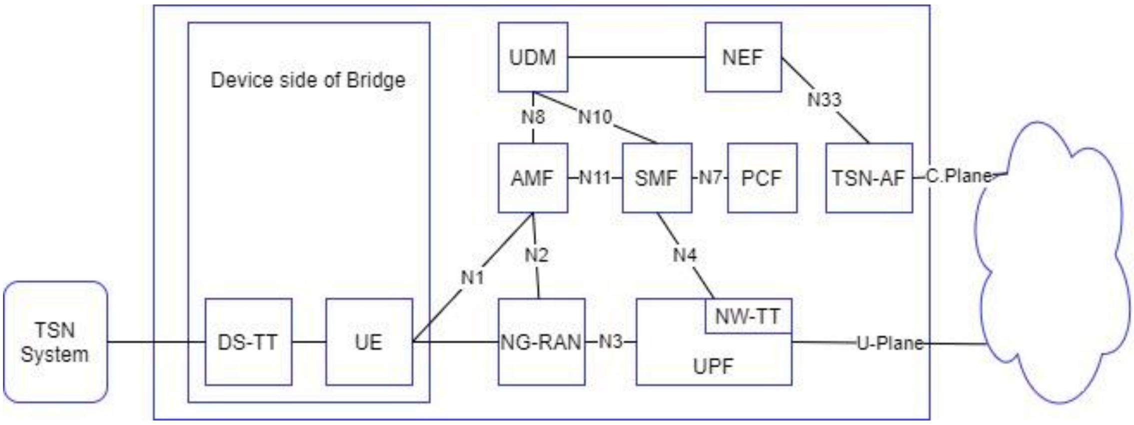


Figure 2.5. 3GPP architecture for TSN integration [34]

Moreover, 5GS supports the creation of group devices as part of the Virtual Network with functions such as 5GLAN and Group Management Functions. The operator of the non-public network will utilize these functions to create groups that will be part of the same LAN including fixed infrastructure and 5GS.

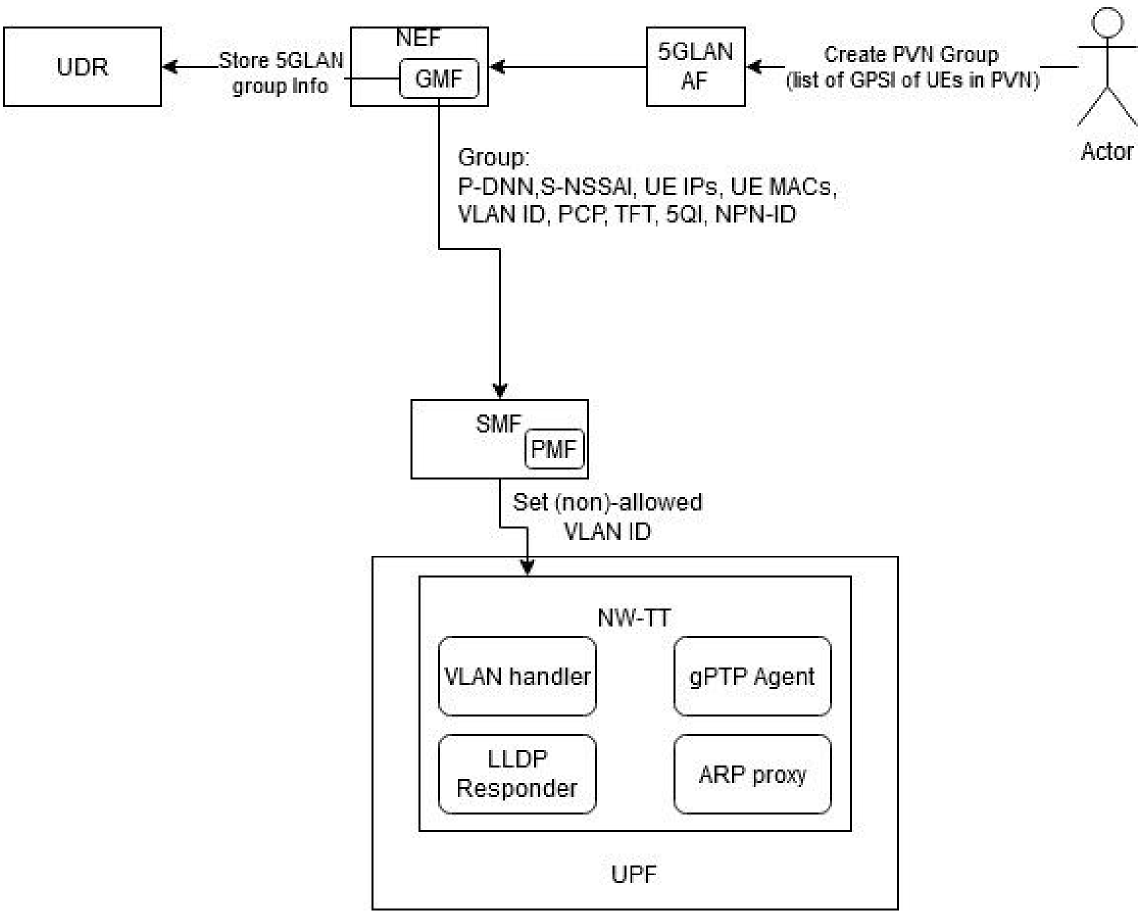


Figure 2.6. 3GPP architecture of Application Functions for 5GLAN and TSN integration

2.2.2.2 NR Reduced Capability

NR Reduced Capability (originally referred as NR-RedCom or NR-Light/NR-Lite) targets three IoT market segments complementary to NR eMBB and URLLC, and also NB-IoT and LTE-M. This NR-based low-complexity solution is

Co-funded by the Horizon 2020
Framework Programme of the European Union

not intended to overlap or replace LPWA use cases based on NB-IoT/LTE-M, but instead, it aims to exploit the potential of NR to serve applications with capabilities ranging from LTE Cat-1 to Cat-4 (i.e. the lowest data rate and bandwidth capacity should be no less than an LTE Cat-1bis modem).

The study on NR-RedCap is addressed from a set of technical features that are compared with a legacy Rel-15/Rel-16 NR device. Legacy NR devices are high-end eMBB reference devices built upon Rel-15/Rel-16 that serve as a basis to analyse the outcomes derived from the NR-RedCap technical reductions.

The key aspects to design a RedCap NR UE are as follows [35]:

- **Device complexity and cost**: NR-RedCap aims to significantly reduce the UE complexity and cost as compared to the legacy NR devices.
- **UE coverage recovery**: Given the degradation resulting from the complexity reduction techniques, NR-RedCap needs to introduce coverage recovery features to mitigate or limit such performance degradation.
- **Device size**: NR-RedCap targets a reduction in the UE form factor as compared to the legacy NR devices.
- **Deployment scenarios**: the system focuses on standalone mode and considers the impacts for a UE that supports multiple RF bands through operation in a single band at a time. It should support all FR1 (410 MHz – 7125 MHz) and FR2 (24250 MHz – 52600 MHz) bands for FDD and TDD. The support of millimetre wave frequencies is a distinctive feature with respect to NB-IoT and LTE-M.
- **UE power saving**: special attention is also paid to reduce the power consumption and enhance the battery lifetime of the devices.

The coexistence of RedCap UE with the so-called legacy Rel-15/16 NR UE needs to be ensured. In this respect, special focus is devoted to specify how the RedCap UE will be explicitly identifiable to networks and networks operators, and how the operators will grant them access or restrict it if desire.

Use Cases

Three use cases are targeted in the NR-RedCap context: (i) industrial wireless sensors, (ii) surveillance cameras, and (iii) wearables. Their requirements are described below (and summarized in Figure 2.7):

- NR-RedCap aims to enable connected industries by means of **industrial wireless sensors** (e.g. motion or pressure sensors, actuators, accelerometers). These sensors must have a small form factor and long-lasting battery life addressing relatively low-end services with requirements higher than LPWA, but lower than eMBB and URLLC. Since URLLC requirements are usually closely related to massive industrial wireless sensor network (IWSN) use cases, a service availability of 99.99% and an end-to-end latency less than 100 ms is expected (indeed, 5-10 ms for safety sensors). Further, the traffic

Co-funded by the Horizon 2020 Framework Programme of the European Union

pattern is potentially asymmetric with a high load in UL, and the reference bit rate is less than 2 Mbps.

- The second use case covers the deployment of **surveillance cameras** used not only in industrial environments, but also in the context of future smart cities. For this purpose, the requirements are: video bitrate should be 2-4 Mbps, latency less than 500 ms, and the reliability 99%-99.9%. A high-end video would require 7.5-25 Mbps.
- NR-RedCap also targets mid-tier **wearables** like watches, eHealth gadgets or monitoring tools whose battery should last several days up to 1-2 weeks. The reference bitrate for these applications are 5-50 Mbps in DL and 2.5 Mbps in UL (peak bitrate up to 150 Mbps in DL and 50 Mbps in UL).

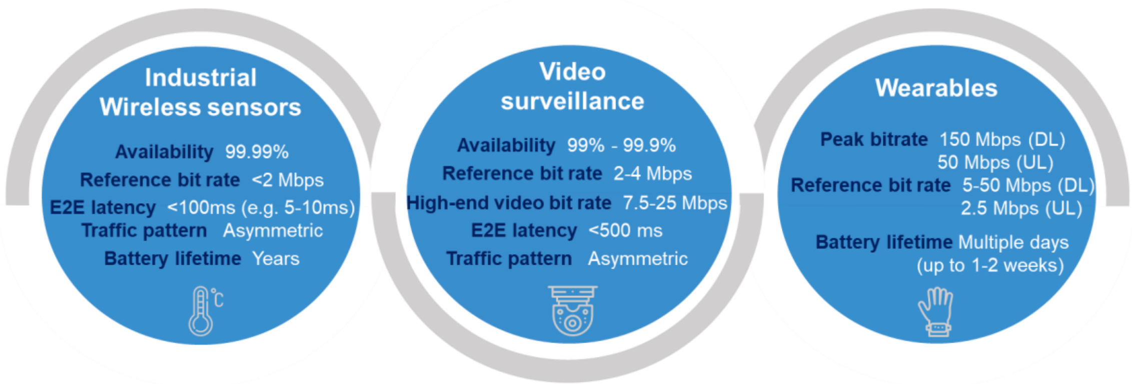


Figure 2.7. NR-RedCap use cases [35]

RedCap NR UE

The main characteristic of a RedCap NR UE is its inherent low-complexity structure compared to a legacy NR UE. Table 3 presents the technical UE complexity reductions specified by the 3GPP in TR 38.875 [35]: (i) reduced number of UE Rx/Tx antennas, (ii) reduced bandwidth, (iii) HD-FDD operation, (iv) relaxed processing time, (v) relaxed maximum number of multiple-input multiple-output (MIMO) layers, and (vi) relaxed modulation order.

Table 3. RedCap NR UE complexity reduction features. Source: [35]

Features	**Reference NR UE**	**RedCap NR UE**
Number of Rx antennas	2Rx (FR1 FDD) 4Rx (FR1 TDD) 2Rx (FR2 TDD)	1Rx (FR1 FDD) 1Rx and 2Rx (FR1 TDD) 1Rx (FR2 TDD)
Max bandwidth	100 MHz (FR1) 200 MHz (FR2)	20 MHz (FR1) 50 or 100 MHz (FR2)
Duplex	FD-FDD	HD-FDD
Max DL MIMO layers	2 MIMO layers (FR1 FDD) 4 MIMO layers (FR1 TDD)	1 MIMO layer (FR1 FDD) 1 and 2 MIMO layers (FR1 TDD)

Co-funded by the Horizon 2020 Framework Programme of the European Union

	2 MIMO layers (FR2)	1 MIMO layer (FR2)
Max modulation order	256 QAM (DL) and 64 QAM (UL) - FR1 64 QAM (DL) and 64 QAM (UL) - FR2	Support of 256 QAM in DL for FR1 optional

Based on the analysis of these reduction techniques, the following is recommended for a RedCap UE:

- One and two **Rx antennas** are supported by the specification. For FR1 FDD and FR2 TDD, the number of Rx branches for a RedCap UE is 1 as compared to a reference NR UE equipped with a minimum of 2 Rx branches. This means that no MIMO capability is offered. For FR1 TDD bands, the minimum number of Rx branches for a RedCap UE is to be selected during the work item phase between 2 or 1 (where 2 would also be supported in the latter case) as compared to the minimum 4 Rx branches for a reference NR UEs. This feature potentially decreases the device size in FR1 (still unclear for FR2).

- The **maximum bandwidth** for a RedCap UE during and after initial access is 20 MHz in FR1 as compared to the 100 MHz bandwidth of a legacy NR device. For FR2, 50 and 100 MHz bandwidth are under study, although 100 MHz is recommended. It is assumed that it applies to both data and control channels as well as both DL and UL.

- NR-RedCap also targets the operation in HD-FDD mode as LTE-based C-IoT. HD-FDD Type B operation, i.e. the larger guard period type, is not supported for RedCap FR1 FDD UEs in Rel-17, so HD-FDD operation Type A and/or FD-FDD are prioritized for FR1 FDD.

- The **maximum modulation order** is also relaxed for a RedCap UE. Support of 256 QAM in downlink for FR1 is optional (instead of mandatory). No other relaxations are supported by specification for a RedCap UE. The study addressed the limitation of the maximum modulation order down to 64 QAM in DL and 16 QAM in UL for FR1 and 16 QAM in DL and UL for FR2.

- The study also considers supporting **relaxed UE processing time** in terms of more relaxed N1 and N2 values (as defined in TS 38.214 [48]) compared to those of UE processing time capability 1. To this end, N1 and N2 are assumed to be doubled. For 15, 30, 60 and 120 kHz SCS, N1 targets 16, 20, 34 and 40 symbols (assuming only front-loaded DMRS) and N2 targets 20, 24, 46, and 72 symbols. This mechanism potentially reduces UE complexity by allowing a longer time for the processing of PDCCH and PDSCH and preparing PUSCH and physical uplink control channel (PUCCH). Relaxed channel state information (CSI) computation time was also considered, assuming doubled Z and Z' compared to the values defined in [48].

The average estimated cost reduction for a NR-Recap UE as compared to a legacy NR device is presented In Appendix A. Both Individual techniques and the combinations of them have been studied. Additionally, Appendix A collects the impacts such UE reduced complexity techniques individually have on the following KPIs: coverage, data rate, latency and reliability,

Co-funded by the Horizon 2020 Framework Programme of the European Union

network capacity and spectral efficiency, power consumption and PDCCH blocking rate.

In addition to the UE complexity reduction techniques presented above, the 3GPP has also studied the following UE power saving mechanisms in [35]: (i) reduced PDCCH monitoring by smaller numbers of blind decodes (BD) and control channel element (CCE) limits, (ii) extended DRX for RRC Inactive and/or Idle mode, and (iii) RRM relaxation for stationary devices. Further details of these UE power saving features are presented In Appendix A. The enhancements introduced by Rel-17 Power Saving work item [49] should also be applicable to RedCap UEs.

Coexistence with legacy UEs

The coexistence of Redcap UE with the so-called legacy Rel-15/16 UE needs to be ensured. In general, RedCap UEs with reduced number of Rx branches can coexist with legacy UEs if early indication of RedCap UE is guaranteed. Otherwise, both legacy UEs and RedCap UEs will be treated the same by the network leading to conservative treatment of all UEs.

The UE bandwidth options also achieve good coexistence performance for 20 MHz and 100 MHz bandwidth in FR1 and FR2, respectively. Indeed, the 20 MHz bandwidth option allows a RedCap UE to reuse existing procedures for acquiring synchronization signal block (SSB), SIB1, other SIBs, Random Access Response (RAR) and Msg4. However, using 50 MHz bandwidth in FR2 would result in a coverage loss for PDCCH reception when control resource set #0 (CORESET) #0 is configured to have 69.12 MHz bandwidth. If early RedCap UE identification is not provided, supporting 50 MHz RedCap UEs requires the gNB to schedule the PDSCH of SIBs, RAR, and Msg4 within 50 MHz bandwidth, causing an impact on legacy UEs.

Further, if RedCap and eMBB UEs share the same initial bandwidth part (BWP) in downlink and uplink for initial access procedure, and the number of RedCap UEs in the network is large, gNB may need to use some means (e.g. access control) to avoid congestion due to high load or configuration restriction, e.g. for random access channel (RACH) occasions.

Introducing HD-FDD operation might make gNB scheduling more complicated. For initial access, supporting HD-FDD Type B operation might have a potential impact on the RACH procedure in that longer time gaps between messages might be needed. This is not an issue for Type A due to its faster UL-to-DL switching capability.

Relaxed UE processing time capability in terms of N1 and N2 may for RedCap UEs increase the complexity for the scheduling. It may cause potential coexistence issues with legacy UEs during initial access if early identification of RedCap UEs prior to Msg2 scheduling is not supported or conservative scheduling is not possible. If gNB schedules all UEs according to relaxed timing relationships for RedCap UEs, legacy UEs may experience an increase in control plane latency.

RedCap UE Types

Explicit definition of RedCap UE type(s) is needed for UE identification and potential constraining the use of reduced capabilities. The capabilities of

Co-funded by the Horizon 2020 Framework Programme of the European Union

RedCap UEs can be categorized as mandatory capabilities that all RedCap UEs support, and possible optional capabilities, signaled explicitly. Rel-16 specifies that the UE reports its radio access capabilities at least when the network requests the UE to do so. The final categorization of capabilities into the studied categories depends on the exact capabilities applicable to RedCap UEs, to be defined during the work item phase.

The number of different UE types should be minimised to reduce market fragmentation. UE types should be introduced only where essential to control UE accesses and differentiate them from other non-RedCap UEs. Having only one RedCap UE type is beneficial for several reasons: (i) there is no market fragmentation of “types”, (ii) a simpler specification (e.g. on early identification, access control, etc.) is achieved and (iii) non-technical discussion outside 3GPP’s scope (e.g. product management, similar to the discussions on LTE categories) is avoided. Otherwise, independent and flexible access control for different UE types cannot be provided.

The study also includes an objective on how to ensure that RedCap UEs are only used for intended use cases, meaning that they only serve services and resources targeted for RedCap UE type. To this end, 3GPP proposes four potential solutions that do not need to be mutually exclusive and are pending to be down-selected in further work item phase:

- RRC Reject based approach: RAN can reject an RRC connection establishment attempt if the service that the UE request is not allowed for RedCap UEs. The service type can be known, for instance, based on the establishment cause provided in Msg3, through higher layer mechanisms or other ways.

- Subscription validation: during the RRC connection setup, the RedCap UE indicates its condition to the core network which validates the UE's indication against its subscription plan and decides whether to accept or reject UE's registration request. For example, network may reject UE if UE indicates it is RedCap, but its subscription does not include any RedCap-specific services.

- Verification of RedCap UE: network performs capability match between UE's reported radio capabilities and the set of capability criteria associated with UE's RedCap type.

- Left up to network implementation to ensure RedCap UE uses intended services and/or resources.

RedCap UE identification and access restrictions

RedCap UEs need to be identified in order to ensure the network can provide services properly in the cell (e.g. scheduling messages and potentially restricting the access to the network). Report time of such identification depends on when the network needs to know the UE type to properly schedule it. Four approaches have been studied:

- **During Msg1 transmission**, for instance, via separate initial UL BWP, separate physical random access channel (PRACH) resource, or PRACH preamble partitioning between RedCap and non-RedCap UEs. Early identification of RedCap UE type(s) during transmission of Msg1

Co-funded by the Horizon 2020 Framework Programme of the European Union

may be necessary for: (i) coverage recovery; (ii) identifying UE minimum processing times capabilities for PDSCH processing and PUSCH preparation; (iii) identifying UE capability for UL modulation order for Msg3 and Msg5 scheduling; and (iv) identifying UE max bandwidth capability for Msg3 and Msg5 scheduling and PUCCH in response to Msg4.

- **During Msg3 transmission**. If early identification of RedCap UE type(s) via Option 1 is not supported, this option may be necessary for coverage recovery for one or more of: Msg4 PDCCH/PDSCH, Msg5 PUSCH and associated PDCCH.
- **Post Msg4 acknowledgment**: for instance, during Msg5 transmission or as part of UE capability reporting.
- **During MsgA transmission**, in case of 2-step RACH procedure: for instance, via separate initial UL BWP, in MsgA preamble part via separate PRACH resource or PRACH preamble partitioning, or in a new indication in MsgA PUSCH part. This option may be necessary for: (i) coverage recovery for MsgA transmission (UE selection of RedCap specific 2-step resources), and (ii) coverage recovery for MsgB and later messages, and associated PDCCH.

Further, NG-RAN supports overload and access control mechanisms aimed at having RedCap specific access restrictions to avoid or limit negative impact on legacy performance. These functionalities are:

- **Cell barring**: an explicit or implicit indication in broadcast system information is recommended to indicate whether a RedCap UE can camp on the cell or not. If a RedCap UE is not allowed to camp on a cell or the RedCap UE considers the cell as barred, it could be of interest to bar all cells on the frequency to ensure RedCap UEs only camp on the strongest cell.
- **Unified Access Control (UAC)**: This framework is specified in [50] and it applies to all UEs in RRC_IDLE, RRC_CONNECTED and RRC_INACTIVE. One option is that UAC can differentiate between RedCap and non-RedCap UEs. The possible solutions for NR-RedCap UAC that have been considered in the study are the following: (i) defining one or more NR-RedCap specific Access Identities, related to the UE type; (ii) defining NR-RedCap specific Access Categories, related to the type of access attempt; (iii) using some of the operator defined Access Categories for RedCap; (iv) broadcasting a different set of UAC parameters for RedCap UEs, allowing network to flexibly and separately provide UAC parameters for RedCap UEs while avoiding impact on UAC configuration of non-RedCap UEs; (v) using existing broadcasted UAC parameters for RedCap UEs with no changes, that is, the same UAC parameters apply for all UEs (non-RedCap UEs and RedCap UEs) and no new Access Categories and Access Identities are defined.
- **RRC connection reject**: Although cell barring and UAC is preferable to save radio resources and limit negative impact on legacy network

Co-funded by the Horizon 2020
Framework Programme of the European Union

performance, it is also possible to use RRC connection reject if the network knows the UE is a RedCap UE.

2.2.2.3 Non-Terrestrial Networks

The growing demand for new services is facing the limitations of current terrestrial systems capabilities. Scenarios that are too expensive or very difficult to cover such as aerial, maritime, or rural need to be addressed, and satellite communications are the only feasible enablers of this connectivity. As this need for NTN appears more evident, several satellite operators have already started the race as service providers. SpaceX is already offering pre-orders for Internet service over their Low Earth Orbit (LEO) constellation Starlink, to be available by the end of 2021. OneWeb is also very close to offering commercial service in rural areas. Other companies such as Telesat or Amazon are also developing their own broadband satellite fleets. Microsoft announced Azure Space to offer modular cloud computing data centres to be deployed over Starlink's and SES' O3b satellites. Lynk is planning to start providing service already in 2022. SpaceMobile already has agreements with Vodafone and AT&T operators to deliver internet connectivity to their customers using their satellites. Sateliot has launched the first of their expected 100 nanosatellite constellation expected to offer IoT services under a 5G architecture from 2022.

In expectancies of this seemingly growing need for satellite communications, the 3GPP already started their work on NTN in Release-15 by trying to specify the deployment scenarios and modelling the new channels. A non-terrestrial network was defined as a network where spaceborne (i.e., Geostationary Equatorial Orbit, Medium Earth Orbit, Low Earth Orbit) or airborne (i.e., Unmanned Aircraft Systems and High-Altitude Platform System) vehicles act either as a relay node or as a base station, thus distinguishing transparent and regenerative satellite architectures. Release-16 work focused on adapting NR to support NTN and further work is being performed in Release-17 to specify the NR features for NTN as well as to study the feasibility of ensuring NTN connectivity of IoT devices in emerging services such as maritime or railway.

Use cases

NTNs can provide wide service coverage by means of spaceborne and airborne platforms that are very resilient to physical attacks and natural disasters. For this reason, the expectancies of NTN will be to upgrade the performance of terrestrial networks in uncovered (remote areas, on board vessels, aircrafts) or underserved (sub-urban/rural) areas, where cost effectiveness does matter; to provide service continuity for M2M/IoT devices, passengers on board moving platforms (ships, bus, high-speed train) or ensuring service availability in critical communications for future railway/maritime/aeronautical communications; and to provide multicast/broadcast resources for data delivery towards the network edges, thus enabling network scalability [51].

The role of NTN in 5G and beyond systems is expected to cover multiple verticals, including transport, Public Safety, eHealth, Agriculture, Finance, Automotive, Media and Entertainment, Energy [57]. In this context, NTN

Co-funded by the Horizon 2020 Framework Programme of the European Union

would support the three main usage scenarios defined by ITU [58]: eMBB, mMTC, and URLLC. However, since critical communications might be hard to accomplish due to propagation delays, the former two use cases will be mainly considered. This is reflected in 3GPP's on-going work presented at the end of this section. eMBB is being considered in an on-going work item to define solutions for NR to support NTN. On the other hand, for mMTC, NB-IoT and LTE-M technologies are being studied for NTN support in a current study item.

Deployment options

Deployment of NTN will be dependent on the type of platform used, which is categorised into spaceborne and airborne. Spaceborne platform category will depend on three main parameters: altitude, beam footprint size and orbit.

3GPP has focused in TR 38.811 [51] on specifying the key features of a few selected deployment scenarios. Mainly, it has focused on Geostationary Equatorial Orbit (GEO) and LEO satellites for 5G NTN deployments, working in two frequency ranges: S-band and Ka-band. GEO satellites have a circular orbit along the equatorial plane at an altitude of 35786 km, with a period of 24 hours, along with earth, giving the impression it is in a fixed position. The beam footprint size ranges from 200 to 3500 km. LEO satellites on the other hand are located at a much lower altitude, between 500 and 2000 km, with circular orbital periods of 94 to 127 minutes. Beam footprint size ranges between 100 and 1000 km.

Airborne platforms being considered are mainly Unmanned Aircraft Systems (UAS), which range from 8 to 50 km in altitude, specifically High-Altitude Platform Systems (HAPS) at 20 km altitude. However, HAPS requirements are considered a subset of those identified for GEO and LEO and so it will be supported by any radio supporting these satellites.

Architecture

The previous section showed a classification of NTN based on the type of platform. However, in terms of NG-RAN, two possible architectures are defined based on the type of the carrier payload: transparent and regenerative.

Transparent or bent pipe refers to the case where only radio frequency filtering, frequency conversion and amplification are performed on the platform whereas the regenerative payload also implements demodulation/decoding, switching and/or routing, coding/modulation; effectively performing all functions of a base station (gNB) on board.

According to this category along with the type of platform, 3GPP defines six possible scenarios. The first two refer to a GEO based non-terrestrial access network for both a transparent and a regenerative payload. The other four consider a LEO based non-terrestrial access network. The differences between the four are, on the one hand, considering either steerable beams or beams that move with the satellite, and on the other hand, whether the satellites use a regenerative or a transparent payload.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Moreover, when considering the NG-RAN, another classification based on the type of access is defined: satellite access architecture and relay-like architecture. In satellite access architecture, the UE connects directly to the satellite, whereas in a relay-like architecture there is an intermediate node acting as a bridge connection between UE and satellite [59]. Figure 2.8 shows the four possible scenarios according to both architecture classifications defined: *payload* and *access*, while interfaces between entities are also denoted.

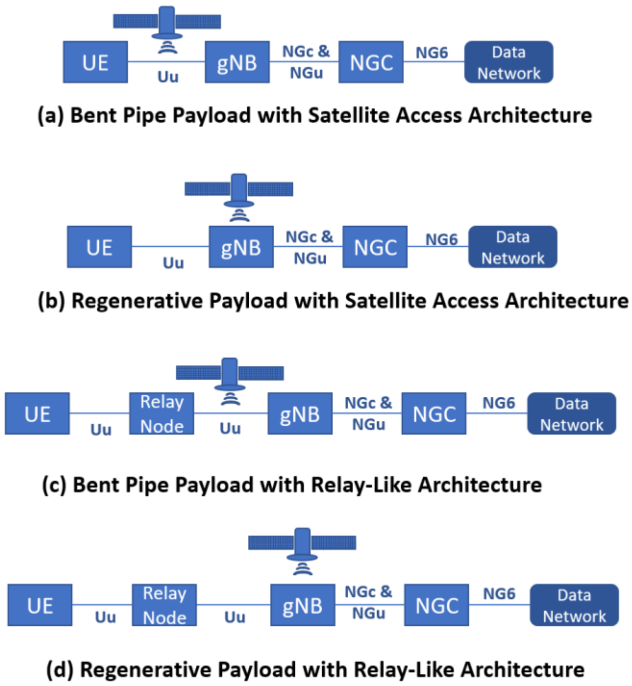


Figure 2.8. NTN architecture options

3GPP on-going work

In Release-17 a new work item is started: *Solutions for NR to support NTN* [60]. The objective of this work item is to specify the enhancements identified for LEO and GEO based NTN with implicit support for HAPS and Air to Ground (ATG) scenarios. Enhancements for NR radio interface as well as NG-RAN are to be specified starting from what has already been studied in Rel-15 and Rel-16. A transparent payload architecture with earth fixed tracking areas and FDD systems is being targeted. Global navigation satellite system (GNSS) capabilities are also assumed for the UEs.

In addition, as we already mentioned in previous section, a new study aiming at evaluating and confirming solutions to address the minimum necessary specifications to adapt NB-IoT and LTE-M for NTN support has also started in Rel-17 [61]. The first objective of this *IoT-NTN* study is to identify scenarios applicable to both technologies assuming that they are working in sub6GHz bands, with either LEO, or GEO constellations, and transparent payload. The second objective is to study and recommend necessary changes to support NB-IoT and LTE-M over satellite for the identified scenarios. The addressed items will cover aspects related to random access procedure and signals,

Co-funded by the Horizon 2020 Framework Programme of the European Union

mechanisms for time and frequency adjustments, HARQ operation, timers, idle and connected mode mobility, system information enhancements and tracking area enhancements. The results of this study item are being documented in TR 36.763 [62].

Appendix B goes into some more detail and analyses the several technical issues discussed currently within the NR-NTN work item, which will also become the basis for the IoT-NTN specification that will eventually follow the current 3GPP study.

2.2.3 Technologies comparison

As we have seen above, several solutions have been introduced over the last years to efficiently connect classical “things” into IoT networks. Depending on the application, the interested user or business can select the most appropriate technology (or set of technologies) to connect objects.

Key advantages of the various non-cellular IoT connectivity solutions (and vice-versa, key limitations of C-IoT connectivity) generally include the alleviation of cost for licensed spectrum resources and the more flexible and simpler deployment to extend existing network coverage to address dedicated IoT use cases.

On the other hand, C-IoT technologies, driven by 3GPP, provide a reliable and secure solution for ubiquitous connectivity of IoT devices and can deliver the following benefits: 1) New IoT services do not require installation of a new connectivity infrastructure to provide communications in the area of the desired IoT service when already covered by a cellular network operator; 2) Reliable and predictable service performance (inherent traits of cellular technologies) while not relying on uncertain availability/interference of unlicensed spectrum, thus, reaching wide coverage quicker and at lower cost from reuse of the existing network infrastructure; 3) Long-term support (since based on global standards) and future proof (since part of cellular communication deployments with plans over decades) with continuous evolution, while keeping backward compatibility for legacy devices.

Table 4 maps the most appropriate C-IoT technology and device solution (“Category”) for typical IoT applications according to their throughput requirements.

Table 4. C-IoT solutions for typical IoT applications according to their throughput requirements

Typical applications	Throughput requirement	Category
Routers network bridges, High-res video, AR/VR devices	1Gbps	**LTE Cat-4 (and higher)**
Video surveillance, In-car hotspot, Infotainment, Digital signage	100 Mbps	

Co-funded by the Horizon 2020 Framework Programme of the European Union

Wearables, Video surveillance	Up to 150 Mbps	**NR RedCap**
Industrial wireless sensors	1-2 Mbps	
Telematics, Predictive maintenance	10 Mbps	**LTE Cat-1 / Cat-1bis**
Smart watches, Point of sales terminals	1 Mbps	
Patient monitor, Alarm panels, Wearables	100 kbps	**LTE Cat-M**
Fitness devices, Trackers, Gas/water meters	10 kbps	
Waste management, Smoke detector	1 kbps	**Cat-NB**
Parking control, Smart agriculture	Few messages	

Of course, apart from throughput, other important KPIs have to be considered to realise the best solution for a specific application and use case as we have already discussed at the start of the chapter and as investigated in several comparisons in literature [63][64][65][66]. Table 5, for example, provides a brief indicative power performance capabilities overview of some low-power solutions. It should be noted, however, that in terms of energy efficiency several communication parameters play role in the end, such as the distances involved, the payload sizes, the coverage conditions, etc. [67].

Table 5. Indicative power consumption performance overview of various low-power technologies

	BT	LoRa	Sigfox	NB-IoT
Tx Current (mA)	12	24-44		74-220
Rx Current (mA)	0.001	12	49	46
Peak Current (mA)	<15	32		120
Max Output Power (mW)	3	25	25	200
Power consumption for UL (mW)		1) 132-475* 2) 625-2250 3) No connection	1) 980-5850 2) 980-5850 3) 980-5850	1) 186-210 2) 362-431 3) 5393-5536
(*) Conditions: 1) Optimal outdoor reception, 2) Medium reception, 3) Difficult indoor reception				

Moreover, Figure 2.9 depicts the capabilities, in terms of bandwidth and nominal range, of the aforementioned C-IoT technologies in addition to non C-IoT technologies (i.e. non-3GPP communication protocols) like LoRa, Sigfox, Zigbee and Bluetooth. Although Wi-Fi and Radio-Frequency Identification (RFID) / Near-Field Communication (NFC) are also included in the graph to provide a more complete context of non C-IoT scenarios, their detailed capabilities are out of the scope of this document. The nominal

Co-funded by the Horizon 2020 Framework Programme of the European Union

range of each individual link can range from a few meters when using RFID/NFC, up to a hundred meters for short-range technologies like Wi-Fi, a few kilometres in urban areas for medium-range scenarios, and over 10 km in rural settings when using LPWAN. The bandwidth on the other hand, which plays key role into achievable throughputs and device complexity, ranges from around 200 kHz in low-cost low-data requirements technologies like NB-IoT and LoRa to 20 MHz and more in technologies that can address heavy traffic requirements such as NR RedCap and NR IIoT.

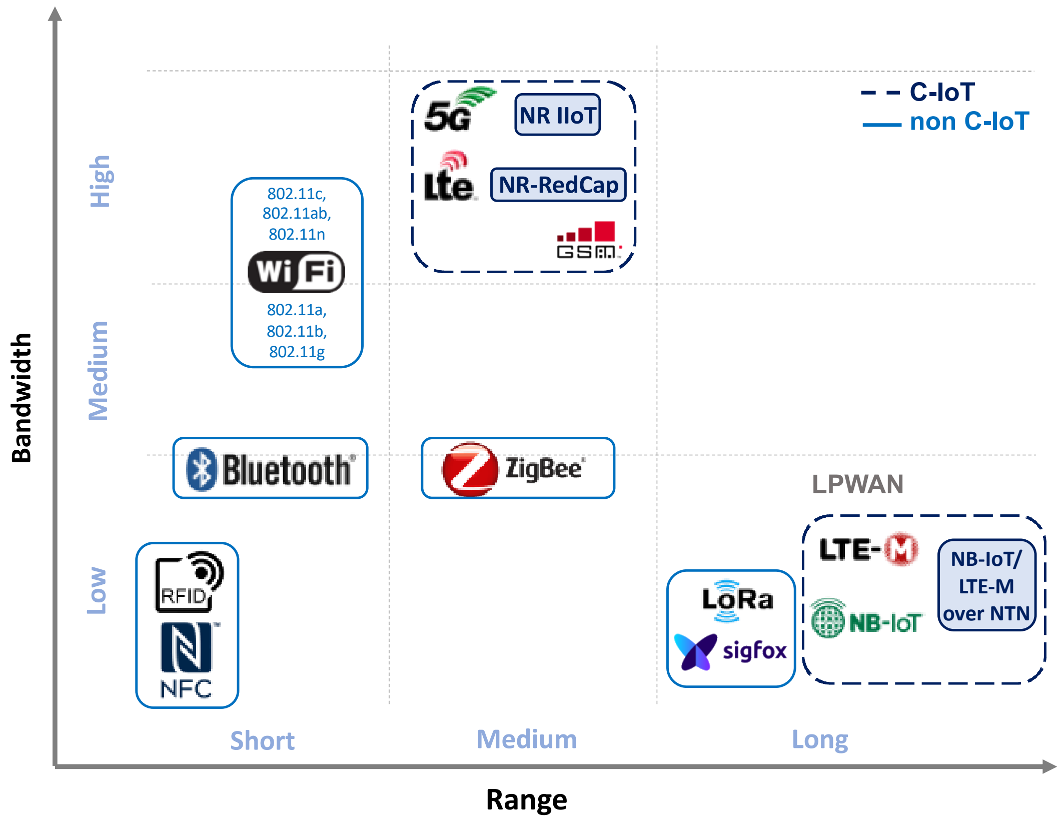


Figure 2.9. C-IoT and non C-IoT technologies

Table 6 below summarizes the bandwidth and range capabilities of various connectivity solutions.

Table 6. Bandwidth and range capabilities of cellular and non-cellular technologies. Sources: [66][67][69]

Capabilities	NFC	BLE	Zigbee	Sigfox	LoRa	NB-IoT	LTE-M
Bandwidth	424 kHz	1-2 MHz	0.3/0.6 – 2 MHz	200 kHz	125 - 500 kHz (UL) 500 kHz (DL)	200kHz	1.4 MHz
Range	< 10 m	10-100 m	10-300 m	10-40 km	2-20 km	> 10 km	

Generally, the range capability, and respectively the topology option for connectivity to the network, is an important aspect. A key advantage of C-IoT versus local connectivity options such as Bluetooth, etc., is that there is no

Co-funded by the Horizon 2020 Framework Programme of the European Union

need to have a local gateway or relay to go to the cloud. For instance, IoT in most common perception current is a jogging watch which is connected to a smartphone; or a temperature sensor which is connected to a home gateway. In these examples, the IoT device sends locally the information to a "box" which then redirects it to the cloud, i.e. there is no direct communication between the IoT device and the cloud. This may be an issue for two main reasons: 1) information from IoT device might be lost if relay is not available; 2) IoT system end to end security is more vulnerable to breach because of this local connection. On the contrary, C-IoT provides a direct connectivity from the device to the network and cloud with higher level of security. Generally IoT devices can connect to the network via three main options: 1) Directly, e.g. using C-IoT or non-cellular IoT technologies, which is more effective for low-power and always-on services; 2) via a terrestrial intermediate device, e.g. gateway, phone/app, etc., for easier integration of heterogeneous solutions; 3) via a non-terrestrial intermediate device, e.g. satellite, HAPS, etc., that can provide ubiquitous coverage even for areas that are remote or not easily accessed by terrestrial network infrastructure. Figure 2.10 illustrates an overview of these communication options and various IoT devices that can be applicable in the selected use cases within iNGENIOUS.

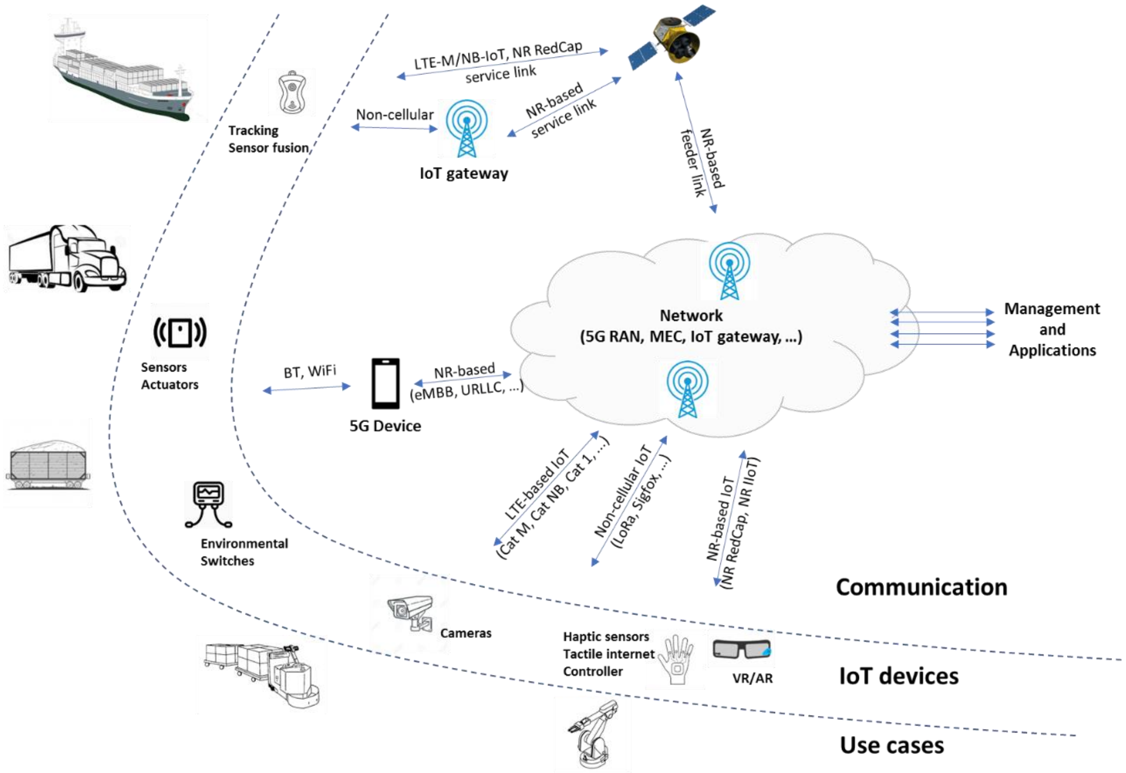


Figure 2.10. Various IoT connectivity options within iNGENIOUS

All in all, there is no universally perfect connectivity solution for all IoT applications; and there are always limitations imposed by any single technology as we have seen in this chapter. Cellular standard offers a global framework in which all IoT applications can find an optimized and secure solution, with global coverage and roaming, from NB-IoT for low end devices, to NR-IIoT for the most demanding one. On the other hand, non-cellular approaches could be appropriate for a local or dedicated deployment with less stringent requirements.

Co-funded by the Horizon 2020
Framework Programme of the European Union

One target of iNGENIOUS will be to identify the appropriate technologies, topologies and device combinations for each one of the use cases under study and focus on improving the minimum requirements for those KPIs that represent a bottleneck in future IoT applications.

2.3 Flexible Software Defined Radio

IoT in iNGENIOUS covers various cases that may require multiple radio solutions, and it is of great interest to investigate SDR-based solution to provide a single hardware/software framework that could be applicable to multiple radio communications standards. Consequently, the design of the air interface's baseband signal at the transmitter is an important aspect for the communications engineers and designers. This can be explored to reduce the impacts of hardware imperfections, such as non-linear power amplifiers and high sensitivity to synchronization errors. Moreover, the structure of the effective channel is also influenced by the baseband design, which impacts the receiver complexity and detection performance [42]. Accordingly, the baseband can be optimized depending on the quality of available hardware and channel conditions. As a reference SDR platform, the National Instruments (NI) Framework provides adaptable PHY and MAC designs, following the IEEE 802.11 and LTE that operates on field-programmable gate array (FPGA) and general-purpose processors. This framework provides an integrated solution with the RF and analog front ends of NI's SDR hardware [43][44]. However, this solution is limited to only commercial standards, and an approach for the design of IoT communications is still missing.

Generally, the connectivity unit at the communication device can be split into three main modules, as illustrated in Figure 2.11: 1) PHY is composed of a baseband module for digital signal processing, and an RF module for analog processing; 2) MAC provides logical channels to carry the payload of the upper layer protocols, and it interacts with the PHY to configure the transmission parameters, such as the carrier frequency, radio resource allocation, and modulation and coding scheme (MCS); 3) the upper layer protocols which work on relatively low-complexity logical operations. The PHY is the critical processing part as it is responsible of the physical transmission and reception, i.e., the translation of binary information into an electromagnetic signal that propagates through the wireless channel.

Co-funded by the Horizon 2020
Framework Programme of the European Union

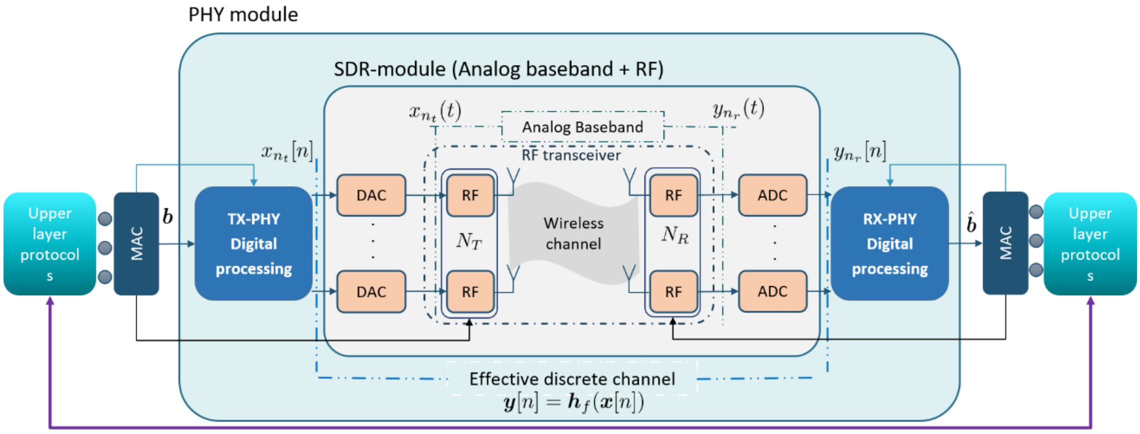


Figure 2.11. Connectivity modules

Most of the RF communications standards employ a similar architecture. The differences lie in the operating carrier frequency and bandwidth, which can be controlled by means of SDR RF frontend, in addition to the baseband and MAC techniques.

As illustrated in Figure 2.11, after the MAC configures the physical transmission parameters, the input data bits $\boldsymbol{b}$ are encoded to multiple digital signals $\{x_{n_t}[n]\}$, where $n_t = 1 \cdots N_T$ is the number of transmitter chains assuming a MIMO system. The digital signals are then converted to analog signals using digital-to-analog converters (DAC) to generate the baseband signals $\{x_{n_t}(t)\}$. These signals are processed by the RF module, which includes low-pass filters (LPF), mixers, and power amplifiers (PA). After propagation through the wireless channel, the RF unit at the receiver, which contains low noise amplifiers (LNA), mixers, and LPFs, outputs the received analog baseband signals $\{y_{n_r}(t)\}$, where $n_r = 1 \cdots N_R$ is the number of receiver chains. After that, the analog-to-digital converters (ADC) convert the analog signals to digital ones, denoted as $\{y_{n_r}[n]\}$. The relation between the transmitted and received signals is defined by the equivalent channel function $h_f(\cdot)$, which elaborates the wireless channel propagation, the hardware response, the additive noise, and interference. The baseband receiver aims at recovering the transmitted data correctly, such that, the estimated data bits $\hat{\boldsymbol{b}}$ are equal to the transmitted data bits with high probability.

Flexible baseband design

In standards like 3GPP 4G and 5G, the baseband unit at the transmitter performs three main tasks, as shown in Figure 2.12. The bit processing block is responsible for performing bit level processing functions including scrambling, channel coding and interleaving. This block can be represented by a function $\boldsymbol{c} = \mu(\boldsymbol{b})$, where $\boldsymbol{c}$ denotes the encoded bits. The encoder introduces additional bits for the forward error correction (FEC) technique, thus, the encoding rate $R_c \leq 1$ defines the ratio between the amount of the information bits and the amount of the encoded bits. The encoded bits are then mapped to digital complex symbols, denoted as $\boldsymbol{d}$, where several bits can be represented by one symbol, such as when M QAM mapping is used. This allows the mapping of $L = \log_2 M$ bits per symbols. Then, the waveform processing generates the discrete IQ sample using mostly linear transforms such as inverse discrete Fourier transform (IDFT) in orthogonal frequency division multiplexing (OFDM).

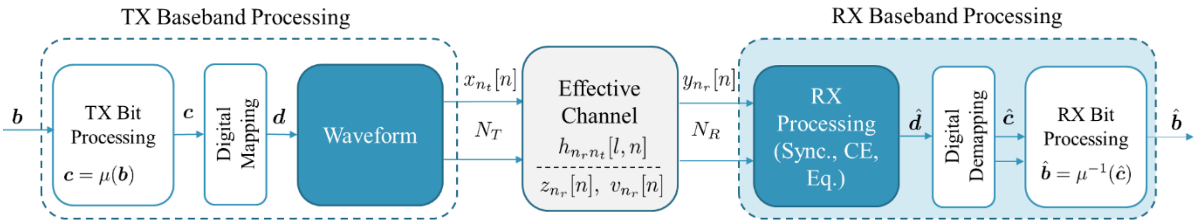
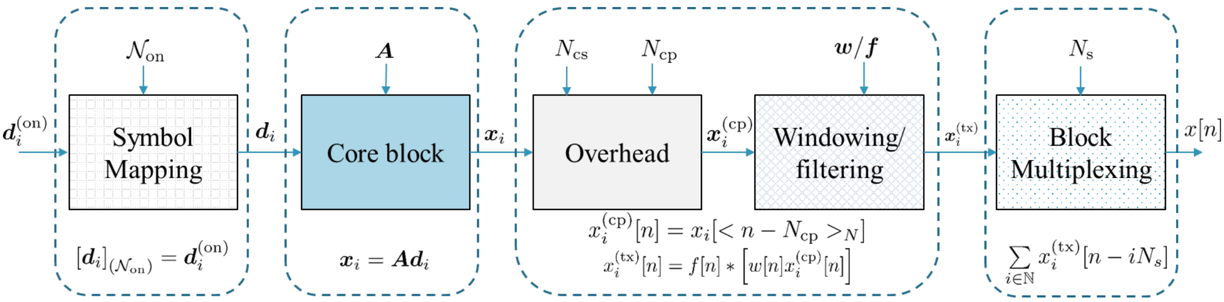


Figure 2.12. Baseband architecture

The baseband processing at the receiver side aims at decoding the information from the received discrete signal. The detection can be implemented in different ways, and it does not need to follow any standard. A common receiver design includes functions, such as synchronization, channel estimation, and equalization, followed by decoding, which can be performed by demapping of the equalized symbols $\hat{\boldsymbol{d}}$, where $\hat{\boldsymbol{c}}$ can be hard or soft bits.

The flexibility of PHY refers to the ability of changing the baseband parameters and functions. Common parameters are the MCS (code rate and the digital mapping order). Additionally, in the 5G NR standard, further flexibility is added by controlling the IDFT transform size. However, the channel coding and digital mapping functions are fixed to predefined implementation. A full flexible solution, able to holistically change the overall baseband function, is still elusive on the current wireless communication standards. This solution allows support to non-OFDM based PHY, and to create optimized functions based on given requirements, hardware constraints, and channel status. Therefore, applications that require specific configurations can be attended by a common hardware architecture and only the software would be modified depending on the application scenario.

A generic linear waveform has been developed in [45] and, as shown in Figure 2.13, the linear waveform can be generated from the data symbols by means of linear transforms using a square complex matrix $\boldsymbol{A}$ of size $N \times N$, followed by additional overhead operations that correspond to adding cyclic prefix (CP) and cyclic suffix (CS). Then, windowing or filtering can be applied to control the spectral shape of the waveform. The block multiplexing is used to realize the linear filtering. The symbol mapping is employed to realize the resource allocation. For example, the modulation matrix is IDFT in OFDM, and the mapping function maps the data symbols to the allocated subcarriers. The flexibility of the waveform is achieved by changing the waveform parameters, i.e. the matrix $\boldsymbol{A}$.



Co-funded by the Horizon 2020 Framework Programme of the European Union

Figure 2.13. Flexible waveform

Alongside the flexible PHY, a flexible MAC is required to convey control information about the used PHY configuration in the case of run-time flexibility. Moreover, a flexible MAC design should be able to work with specific compile-time configurations; for instance, to realize a standard MAC with a standard PHY. A customized MAC design should also be considered based on the use case. For example, a simple MAC that distributes with deterministic resource allocation can be considered to connect few devices to an access point. This helps avoiding the complexity and redundancy of standard MACs that are intended to support general use cases. The flexible PHY/MAC at the UE side is supported by flexible PHY/MAC at the access point, which is reported in D4.1 [3].

2.4 iNGENIOUS Innovation

For the "*things*" to become IoT devices and communicate data, it is required that they include embedded computing units as well as network communication hardware and software to be connected to an IoT network.

Accordingly, one of the aspects iNGENIOUS aims to evolve for IoT devices layer is their *connectivity* solutions in order to leverage their capabilities, address their current limitations, and ease their adoption in the selected use cases. Two main directions of evolution are considered in iNGENIOUS:

- Enhancements to lower the cost of communication (e.g. in terms of computational complexity, power consumption, latency, flexibility), leveraging improvements discussed in the standardisation bodies and investigating new innovative solutions.
- Flexible, software-defined PHY/MAC with different flexibility level at compile-time and run-time for customization and optimization to specific IoT air interface as well as to support different traffic classes (e.g. eMBB, URLLC, mMTC).

2.4.1 C-IoT air interface enhancements

As shown in Section 2.2, 3GPP is continuously improving the support of C-IoT in its standards. LTE-M and NB-IoT are the solutions currently on the market to fulfil the needs of the IoT ecosystem, but NR already included in Release 15 new features to support IoT that were later enhanced in Release 16. The current work in Release 17 is focused on NR-RedCap and on an enhancement of the support of NR IIoT and for NTN.

From this whole set of potential IoT-related improvements being discussed in the 3GPP, we will analyse the potential of a subset of features to provide benefits in some use cases considered in iNGENIOUS. For example, focusing on Industrial IoT, as we have seen in Section 2.2.2.1, RAN-related features under study in Release 17 are: i) physical layer feedback enhancements, ii) uplink enhancements for URLLC in unlicensed controlled environments, iii) intra-UE multiplexing and prioritization of traffic with different priority, iv) enhancements for support of time synchronization, and v) RAN enhancements based on new QoS related parameters (refer to [36] for a

Co-funded by the Horizon 2020 Framework Programme of the European Union

work item description on these features). In addition, there is another ongoing work item focused on IIoT related to 5G-system architecture aspects (see [37] for a description of this work item). In that work item the objectives are i) uplink time synchronization, ii) UE-UE time sensitive communications, iii) exposure of QoS and related parameters, iv) exposure of time synchronization and support for point-to-point time synchronization, and v) use of survival time for deterministic applications in the 5G-system. Thus, concerning these IIoT related features, an initially selected subset of features within iNGENIOUS includes: i) the enhancements of URLLC in unlicensed environments, ii) the enhancements for time synchronization (both at RAN and System levels), and iii) the enhancements based on new QoS parameters and its exposure in the 5G-system. In parallel, the NR-RedCap work (for Factory UC, which considers the use of devices fitting RedCap UE capabilities) and the NTN works (for Transport and Ship UCs, which consider the involvement of satellite-based communication) will be also considered and similar feature shortlisting is undergoing.

System and Link level Simulations

One of iNGENIOUS aims is to compare the current system level performance of available solutions with the performance of the novel proposed solutions that could be available with Release 17 and beyond. To achieve this goal, we will rely on system and link level simulations.

The iNGENIOUS project validates the project outcomes mainly in four (4) large-scale Proof of Concept demonstrations. This strategy will be complemented by simulations of the communications networks used in iNGENIOUS conducted in specific scenarios. Simulations allow us to foresee the network performance before the actual deployment. The drawback is that simulation is not reality and the simulations models used must be carefully selected so as to achieve enough accuracy in the simulation results, close to those that would be obtained from real-life trials. Trials have usually high costs, are limited in human, physical and time resources, and hence provide a very valuable but limited amount of information. On the other hand, simulation is an effective and efficient way of executing a high amount of different test variants, enabling the gathering of huge amount of data. For example, in a simulation, it is easy to consider enormous densities of devices while in a real test the number of connected IoT devices can be much lower. Another great advantage of simulation is that technologies not yet available in the market can also be evaluated with a simulator. For example, in iNGENIOUS project, NR-RedCap is planned to be evaluated thanks to simulations.

iNGENIOUS current strategy for the simulation campaign will be to firstly evaluate the 5G technology and, once proper results are extracted, to move in studying optimization of NR-RedCap, obtaining results for the new 5G reduced version. Simulations campaigns will be carried out mainly in two use cases: "Automated robots with heterogeneous networks" (Factory UC) and "Improved driver’s safety with mixed reality and haptic solutions" (AGV UC).

As for the actual simulator to be used within iNGENIOUS, the basis will be the NS3 system simulator [38]. Specifically, to simulate 5G networks, a 5G-NR module of NS3 called 5G-LENA [39] will be employed. For the link level results, which characterize the physical layer of the 5G system, we will complement the models available in 5G-LENA with new models that consider the newest user equipment devices defined by the 3GPP.

iNGENIOUS contribution will focus on different activities:

- Link Level simulations to complement the 5G-NR physical layer abstraction of the NS3 code. The system simulator includes default lookup tables, but these tables will need to updated with the latest 3GPP MCS configurations regarding robust MCSs.
- Identify the most promising MAC and PHY features to enhance the performance of the 5G system focusing on the IoT-related scenarios of the use cases under study. NR-RedCap will be considered.
- Simulate the operation of the 5G network in specific realizations of the use cases under study, focusing on the most promising MAC and PHY features previously identified.

Simulations in the Factory UC

This use case is based on an indoor industrial scenario, more precisely on the ASTI factory, where the trials will be carried out along the project. This kind of indoor scenarios is characterized by a high density of blockers, especially metallic elements (machinery, assembly lines, shelves, etc.), not found in other indoor scenarios such as the offices. Therefore, the radio propagation is also very characteristic.

Given the increasing interest in the deployment of IIoT, 3GPP decided to develop a new channel model for IIoT, different to the Indoor Hotspot previously used for any indoor evaluation. As a result, 3GPP has recently defined specific scenario within its channel model in TR 38.901 [40], known as Indoor Factory (InF), to consider the indoor industrial scenarios. This model is based on a set of basic assumptions on: antenna height and configurations, room size, UE distribution and movement, clutter type and density or frequency bands, among others. A summary of the model development process and a summary of the 3GPP document can be found in [41].

In iNGENIOUS, we plan to study the applicability of the 3GPP InF channel models in the ASTI factory scenario. Specifically, we will determine the proper parameterization of the model based on above mentioned parameters such as the clutter type. Note that the channel model parameters are not the same; for example, in mixed production areas with open spaces and storage/commissioning areas and in areas with assembly and production lines surrounded by mixed small-sized machineries.

In this use case, several KPIs are intended to be evaluated via simulation. These KPIs are a subset of the whole list identified in D2.1 [1]. Specifically, we will consider: mobility, throughput, end-to-end (E2E) latency, availability, and reliability.

Simulations in the AGV UC

Co-funded by the Horizon 2020 Framework Programme of the European Union

In this use case, there will be two different scenarios, both in the port of Valencia: one corresponding to an indoor spot, and the other to an outdoor one. For these simulations, standard 3GPP scenarios (like indoor hotspot or indoor office) will be employed [40]. Once the locations of the trials are clear, the proper modifications to these 3GPP scenarios will be made for a better accuracy of the simulation results.

The KPIs considered to be evaluated via simulation in this use case are the same presented for the previous use case: mobility, throughput, E2E latency, availability, and reliability.

2.4.2 Flexible HW/SW architecture

The concepts present in this subsection will be key factors for overcoming the challenges listed in D2.1 (Use cases, KPIs and Requirements) [1], since the heterogeneous nature of the network requires a flexible approach not only on higher layers, but also at the lower MAC and PHY layers.

Standard PHY/MAC design is essential for commercial purposes. However, a device with fixed PHY/MAC implementation on dedicated chip only operates with that standard. Flexible implementation allows devices to be upgradable and reconfigurable to extend their life cycle. Moreover, the flexibility enables the development of private PHY/MAC in a private network, such as industrial networks. Finally, for research purposes, flexible PHY/MAC allows real-time experiments of new innovation in PHY/MAC design under realistic channel conditions instead of simplified simulations.

The baseband hardware processing requirements depend mainly on the required data rate, the regularity of transmissions, and the latency. For low data rate and occasional transmission, it is feasible to realize the PHY on general purpose central processing unit (CPU), for some other case a digital signal processor (DSP) is sufficient, and in other situations, hardware implementation is unavoidable. By using reconfigurable hardware architecture, such as FPGA, both types of flexibility are beneficial to optimize the hardware resource consumption at the compile time, and to allow significant flexibility in run time by proper design of the architecture.

A compromise between software and hardware implementation can be exploited with the *kernel-based architecture*, as shown in Figure 2.14. In this architecture, CPUs are used to perform light computation and data routing, whereas the heavy computation is performed on dedicated hardware accelerator processing units (PU) with sufficient run-time flexibility. For example, a PU can be used for the realization of flexible discrete Fourier transform (DFT). In addition, a unified computation architecture can be exploited to run the communication stack and further processing concerning the applications. In particular, each layer can be executed on a dedicated CPU, and the hardware accelerators are shared and coordinated by the scheduling unit.

Specifically, within iNGENIOUS Factory UC, a low latency flexible PHY/MAC will be designed to operate on a tile-based hardware platform, i.e., a hardware-software co-design shall be employed to optimize resource usage, or fit application specific requirements. Thus, security and low latency,

Co-funded by the Horizon 2020 Framework Programme of the European Union

among other performance metrics, can be obtained from a tailored architecture depending on the application scenario. This design approach will be fundamental for developing a secure yet energy efficient IoT UE, as well as the performance oriented gateway. Furthermore, different wireless communication standards can be implemented by simple software updates, without the need for hardware replacement. Flexible HW/SW architecture can also be applicable to the microkernel-based architecture that will be described in Section 3.2, thus, the concept may also be investigated further within the Transport UC.

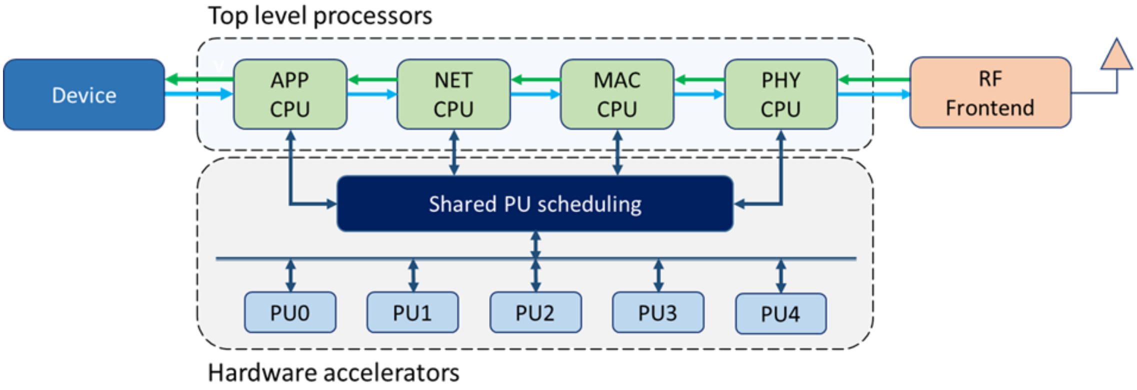


Figure 2.14. Flexible HW/SW architecture

2.4.3 Versatile 5G modem solution

5G modems and routers are being used to connect specific vertical components, such as robots, sensors, cameras, or AGVs to the 5G network. These communication devices are physically connected to the end devices and permit them to communicate with the Next Generation NodeB (gNB) wirelessly, enabling the machines to "speak" 5G. Figure 2.15 shows a schematic of the 5G architecture in a simple way (where Uu and NG-c/NG-u denote the different interfaces). As shown in the figure, the modems or routers could be integrated within the end devices or could be connected to them externally.

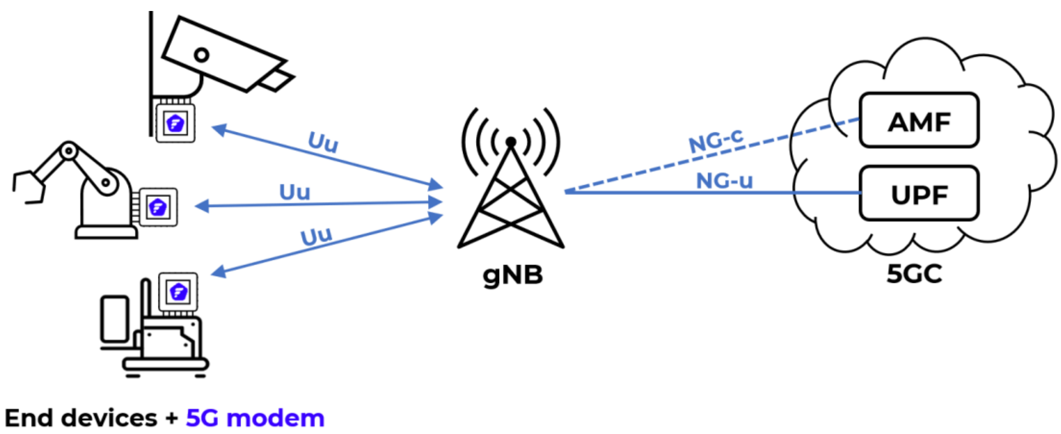


Figure 2.15. Connection between the 5G communication device and the rest of the network components

For a proper implementation for enabling IoT, the 5G communication device must be compact and integrated within the end device, power efficient, simple (used as 'plug and play' device), NSA (non-standalone) and SA (standalone) compatible. Note that an SA deployment here refers to the use of a single cellular technology, namely 5G, in both RAN and core parts, while NSA consists of the combination of 4G and 5G components in the same network.

iNGENIOUS plans to use the Fivecomm 5G Modem (F5GM) to develop, integrate and validate (in Factory and AGV UCs) a compact and flexible end device solution that provides 5G wireless connectivity and that can be customized. F5GM is a board that integrates a third party 5G module and multiple communication interfaces (Ethernet, digital/analog ports, etc.). The board can be protected with a case, being connected with the end devices with a wire. In a typical scenario, the F5GM connects via Ethernet to the end device and connects to the 5G network via its integrated or external antennas. In addition, the F5GM board can be integrated into the end device. The objective is to develop a solution particularized for the connection link between the end device and the 5G network, depending on the specific needs of the use case.

F5GM is a powerful, versatile, and compact device designed to bring all the advantages of the new 5G technology. The F5GM has simplified its electronics to make the most of the 5G modules in the market while minimizing the power consumption and cost. The F5GM is shown in Figure 2.16.

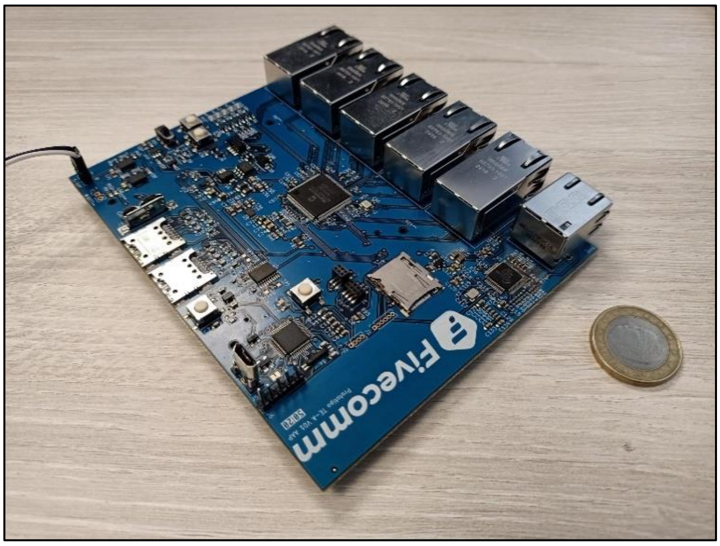


Figure 2.16. Fivecomm 5G modem (F5GM) prototype

Regarding innovation, the 5G modem is expected to implement the following functionalities:

- **Easy deployment**: in a 'plug and play' fashion. It only needs to connect the Ethernet port to the end device, fix the device to the infrastructure and press the 'ON' button.

Co-funded by the Horizon 2020 Framework Programme of the European Union

- **Customization**: Different IP protection degrees are available, with up to 4 internal antennas or external ports to provide the best experience even in low coverage scenarios.
- **Remote management**: It will include a management platform that allows to configure, monitor, and perform software updates remotely.

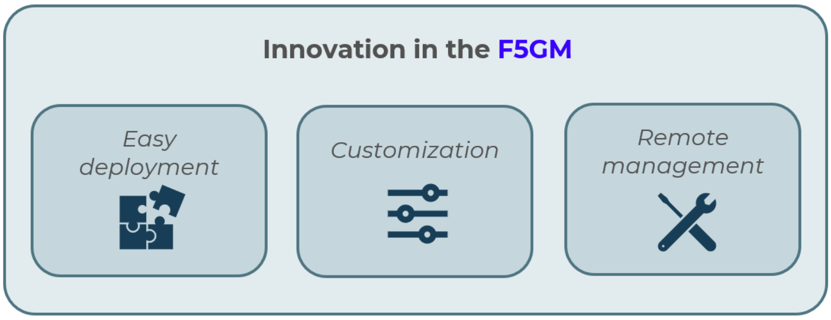


Figure 2.17. Innovation in the Fivecomm 5G module (F5GM)

2.4.4 Relation to UCs

This section refers to the relationship of the state-of-the-art air interface technologies and expected innovations mentioned above with each use case in iNGENIOUS. Table 7 below lists the main IoT connectivity functionalities introduced in the previous sections and, for each of them, how they map with the various relevant iNGENIOUS UCs in terms of innovation maturity. For this mapping, three options are considered:

- **State-of-the-art (SoA)**, meaning that the related functionality is implemented with state-of-the-art technologies.
- **Innovation concept**, for those innovative functionalities, to be investigated within WP3, which are relevant to the UCs but will not be demonstrated.
- **Innovation demonstrated**, for those innovative functionalities, to be investigated within WP3, which will also be implemented and showcased as part of the UC.

The mapping provided in Table 7 is to be considered valid at the time of writing, and it may evolve in the future according to the evolution of UCs implementation and requirements in terms of devices and air interface connectivity. Further detail of each use case description can be found in D2.1 [1].

Table 7: NG-IoT connectivity functionalities mapped to UCs

Use Cases → Functionality ↓	**Factory**	**Transport**	**Port Entrance**	**AGV**	**Ship**	**DLT**

Co-funded by the Horizon 2020 Framework Programme of the European Union

Non-Cellular IoT	SoA	SoA			SoA	SoA
LTE-based IoT	SoA	SoA	SoA		SoA	SoA
NR-based IoT	SoA			SoA	SoA	
NR-RedCap	SoA					
NR/IoT-NTN		Innovative Concept			Innovative Concept	
Flexible HW/SW architecture	Innovative Demo	InnovativeConcept				
Versatile 5G modem	Innovative Demo			Innovative Demo		

Co-funded by the Horizon 2020 Framework Programme of the European Union

3 Ultra-safe Low-Power Dedicated Platforms

The Internet of Things is already widely deployed and in use, but some aspects especially regarding mobile devices have a lot of potential for improvement. The challenges lie in two main areas:

1. Optimizing power consumption to maximize the time a device can remain active in the field before its end-of-life or between service intervals
2. Harden the device against malicious attacks to keep the IoT as a whole trustworthy

The first challenge is typically related to power consumption and limited energy resources. The solution is typically related to power-optimized transmission networks, reduced communication payload, and optimized edge-computing structures. In a nutshell, one should: i) use the least power consuming communication network available for “Alive” information, ii) use the most power optimized communication network for the required payload size, and iii) reduce the edge computing power by lowering clock rates and, if possible, transistor count. LoRaWAN and Bluetooth are an example of relatively low power for near and very near field communications, respectively. And neuromorphic networks are examples of highly optimized parallel computing for real-time and low power computing with reduced transistor count. Both will be discussed in the following Section 3.1.

Addressing the second challenge requires careful design of both hardware and software. Isolation-by-default architectures in combination with a hardware root-of-trust can significantly increase the level of security, reliability, and trustworthiness. This aspect will be discussed in Sections 3.2 and 3.3.

In Section 3.4, the innovations and improvements over the state of the art that iNGENIOUS partners contribute will be discussed.

3.1 Edge IoT Sensor Solutions

The basic challenge of Mobile Edge Connectivity is illustrated by the following paradigm:

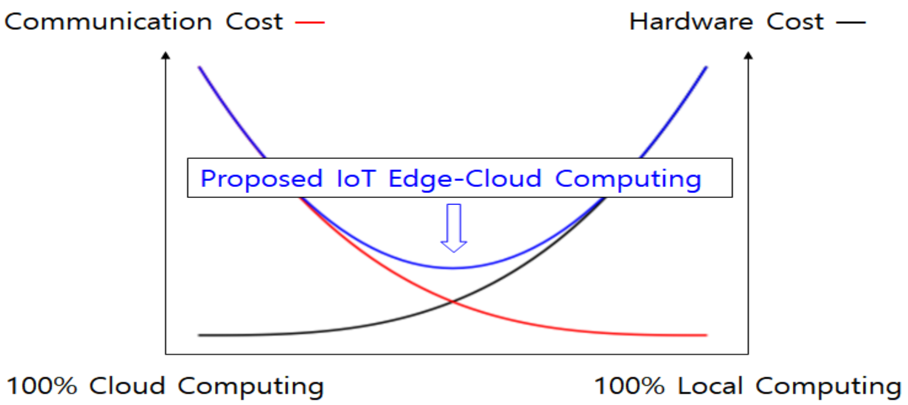


Figure 3.1. Edge vs. Cloud Computing

Communication cost consists of transmission and computing costs. Transmission costs are driven by payload volume and energy requirements to communicate the payload. Computing costs, on the other hand, are driven by hardware costs and energy requirements to reduce communication payloads. In energy-limited applications, payload reduction (raw data to meta-data conversion) is typically the more energy-efficient approach.

To summarize, the balance between cloud and edge computing is driven by the availability of connectivity, connectivity cost, acceptable information latency, and available energy. The optimal solution depends on the use case. There is not one solution but rather a suitcase of solutions which must be optimized to specific requirements. For further discussion on this optimization problem and trade-offs, please refer to iNGENIOUS D2.1, Section 5.2 [1].

The iNGENIOUS approach specifically focuses on energy-limited applications which are typically mobile or battery-operated applications. The payload of these applications varies. This variation can be due to dynamic shifts in edge vs. cloud computing, or because of dynamic resolution requirements, or learning strategies. The state machine implementing such a hybrid approach is application specific. In general, however, the communication approach should be suitable for the payload which is to be transmitted. Without being all inclusive, the overall communication approach (from sensor generated data to air interface) is summarized in Figure 3.2. This will also be the approach that will be used for iNGENIOUS transport UC, where examples of low-mid-high volume data transmissions (indicated by the green blocks) will be demonstrated, while miniaturized IoT system (MIOTY, see below) and NR-based functionalities will not be demonstrated (indicated by the orange blocks).

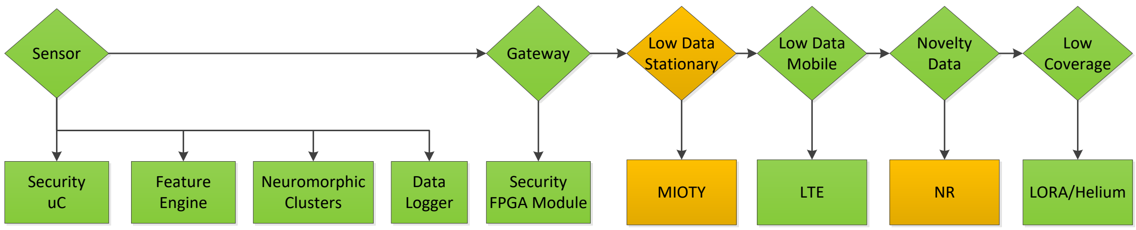


Figure 3.2. Situation-based energy optimized edge computing

Let's assume the edge device is capable of Raw to Meta-Data reduction. We simply assume that processing energy for data reduction is a factor of 10-100000 times less than communication energy of unprocessed data. (Example: 20s or 48 kHz 16bit Data = 1.92 MBytes vs. 16 variable feature vector (i.e. Meta-Data) at 16bit = 32 Byte = Factor 60k). The task of the edge sensor is to minimize or avoid raw data transmission whenever possible. Thus, the following general steps can be considered:

- Whenever possible, the edge sensor shall reduce raw data to signal features.

- Whenever possible, signal features shall be classified into relevant signal states.
- If required, the edge sensor shall differentiate known from unknown states.
- If required, the edge sensor shall cluster, and count known and unknown states.
- If required, the edge sensor shall save selective data on unknown states for cloud machine learning (ML).

In the end, what information is transferred and how it is transferred depends on the application:

- For stationary low power applications, MIOTY is a new form of scheduled low-power communication between sensors and gateways, with similar performance as LoRaWAN, but increased robustness in information crowded environments with lots of interference (see Figure 3.3). By using smaller redundant data packages and time scheduled communication slots, MIOTY reduces uncontrolled cross-chat, while allowing sporadic interference without interrupting overall data transmission. MIOTY can also be used in mobile environments such as ships and trains, when Bluetooth connectivity between Edge Sensors and Edge Gateways requires increased communication range.

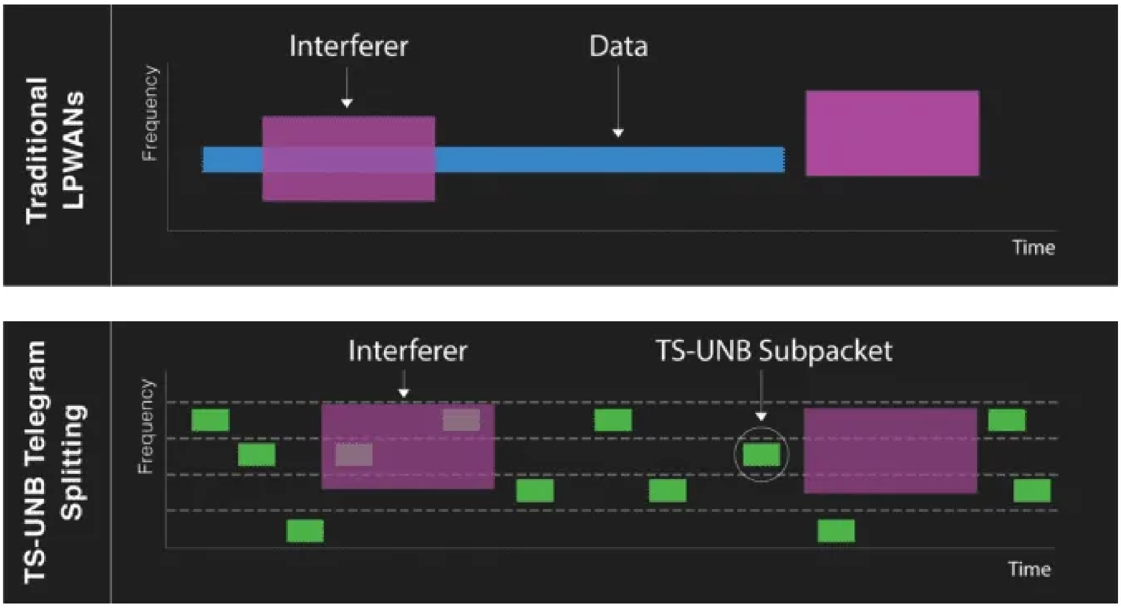


Figure 3.3. Top: Typical COM / Bottom: MIOTY redundant subpackets for communication in interference rich environments

- For communication of small package sizes over larger distances, LTE-based communication is fairly well established. This is typically suitable for metadata or vectorised signal features for cloud-based trend analysis and classification.

Co-funded by the Horizon 2020 Framework Programme of the European Union

- If raw data is to be extracted from edge sensors that have identified and stored novelty clusters, then high bandwidth communication such as 5G is suitable for mobile applications.
- In case mobile networks are not available, then custom mesh networks based on LoRaWAN with extended mobile or satellite node connectivity are an option for cloud connectivity.

The iNGENIOUS Transport UC targets to demonstrate Low-Mid-High volume communication concepts as illustrated in Figure 3.2. However, this use case goes even further. While optimizing communication energy requirements based on payload is by far the most important factor for ultra-safe low-power dedicated edge solutions, minimizing the processing energy of edge computing to achieve data reduction or Meta-Language communication is also an important factor.

A starting point for edge computing power optimization is reducing the clock rate to the lowest rate capable of achieving real-time computation. Interim storage of large data in static random-access memory (SRAM) prior to pre-processing is to be avoided. Likewise, interrupt handling can greatly affect power consumption. Reduced instructions set computers (RISC) - such as ARM processors - are the de facto standard for low-power computing. Smaller silicon pitch size further reduces energy consumption. 3nm technology as the next step in that direction is being worked on. Reducing computing power consumption even further could be achieved by reducing transistor count. An interesting example of transistor count reduction is for example dynamic vision cameras. Rather than sampling real optical images, dynamic vision cameras sample net changes between adjacent frames. The effect is a net-motion image as shown in Figure 3.4. In many applications, such net images greatly reduce data volumes and application specific computing power [51].

Figure 3.4. Data reduction via dynamic vision – net-motion imaging

Neuromorphic computing is another approach to greatly reduce computing power. Even RISC processors have a huge computing overhead. A processor is a general-purpose computation engine and, being man-made, the underlying processing architectures are not biology inspired. Biological processing power consists of neurons which are highly optimized structures to recognize specific patterns. A matching pattern triggers a moderate to large response. Neuromorphic computing follows this approach. A neuromorphic computing network is a massively parallel network of highly specialized pattern engines which spikes above predefined thresholds. A classification can be based on the summary of all spikes or on winners (i.e. winners takes all). More importantly however, and that differentiates neuromorphic networks from neural networks, neuromorphic computing is not about best matches, but best-known matches. In other words, neuromorphic networks don't cover the full feature space, just a portion, and are therefore ideally suited for not just identifying matches but also novelties (see Figure 3.5 for an illustration of this clustering concept).

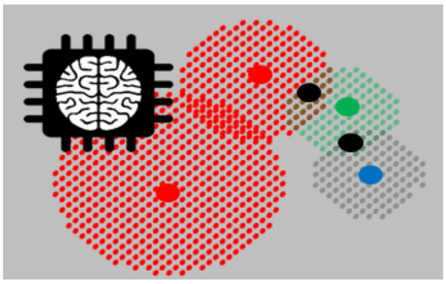


Figure 3.5. Neuromorphic Clustering (Known Clusters & Unoccupied Gray Feature Space)

A neuromorphic neuron is energy efficient, because clustering requires a relatively primitive pattern algorithm with greatly reduced number of transistor count compared to a general processing unit.

Unfortunately, neuromorphic algorithms are not generic, and neuromorphic chip production, if available at all, is typically not produced in latest production technologies. For small volume productions old nodes like 130nm are used. Some newer developments have made it to 55nm production technology. This however cannot be compared with modern micro-controller production technologies which are at 7nm. Therefore, the theoretical energy savings of neuromorphic networks given by their much-reduced transistor count cannot be achieved.

The iNGENIOUS Transport UC uses neuromorphic networks not for operational power reduction but for efficient data acquisition. Neuromorphic networks are massively parallel. This makes them extremely attractive for streaming data analysis. In the use case, iNGENIOUS utilizes neuromorphic networks for data statistics on known clusters and novelty detection. This approach revolutionizes data engineering which is the basis of data-science. Typically, data scientists have either too little or too much data, but seldomly unbiased balanced data. Transport UC leverages neuromorphic cluster networks for long-term studies to measure and collect context-specific cluster space of a given application. Instead of native neuromorphic silicon, FPGA-based neuromorphic networks are used. FPGAs are not energy optimized, but they are extremely fast to analyse massive data streams in real-time. And compared to modern GPU architectures, FPGA engines are comparatively low power.

Neuromorphic Computing Definition

- Neuromorphic Computing is essentially parallel computing with optimized computing engines (Cells)
- Neuromorphic Cells don't have a processing overhead, therefore they are energy efficient

Co-funded by the Horizon 2020 Framework Programme of the European Union

- Parallel computing is needed if the conventional processing time is not enough within the allocated processing cycle
- FPGAs are ideally suited for implementing Neuromorphic computing approaches

Ideal Use Cases of Neuromorphic Computing

- Continuous Data Stream Analysis – Cyber Attack
- Real Time Data Stream Analysis – Data Labelling 🡐 iNGENIOUS Transport UC

3.2 Securing Embedded Computers in IoT Devices

The IoT is a distributed system, in which edge devices such as sensors and actuators communicate directly with each other, with a nearby multi-access edge computing (MEC) server, or servers in a remote data centre. To enable this kind of connectivity, IoT devices require local computation capabilities to run network protocols, as well as cryptographic algorithms that protect the contents of the communication. As a result, IoT devices run large amounts of complex software in order to meet these functional requirements. So, on one hand, this complexity is essential for the device to fulfil its purpose. But on the other hand, complexity is the enemy of security and reliability. And yet, the IoT as a whole can only be trustworthy, if the edge devices can be operated securely and reliably. The risks that are inherent to the functional complexity must be minimized and the internal structure of both the hardware and the software of IoT devices is critical to reaching that goal.

3.2.1 Secure-by-default operating system architecture

On the software side, IoT devices often run a monolithic OS, in which all functionality is part of a single, potentially huge code base with no isolation between subsystems. This is true both for devices with low-power processors that only support a real-time operating system (RTOS), and for those capable of running the much more powerful Linux-based OS. The former type of OS supports only weak isolation or no isolation at all between the OS and device-specific software running on top. The latter requires an OS that is much bigger and more complex to begin with. But here, too, system designers often add device-specific software with little regard to security, as they chose the path of least resistance and allow all device-specific programs to run with unrestricted privileges. Monolithic, high-privilege codebases are a security hazard, as an attacker who manages to exploit a single programming or configuration error anywhere in the codebase will be able to compromise the entire software stack running on the device.

From the security point of view, the state-of-the-art in system-software design is microkernel-based operating systems [52]. In contrast to monolithic systems, the OS itself is split into separate components that run isolated from each other. Only those components that need to cooperate are allowed to communicate with each other, whereas unrelated parts of the system (e.g., from another subsystem of the OS) cannot be accessed at all; the microkernel enforces these restrictions. A security vulnerability in one of the isolated components (e.g., the network device driver) may only lead to a compromise of this one component, but not the rest of the system. Thus, the attack surface of the overall codebase is greatly reduced by this *isolation-by-default* design. In fact, the underlying construction principle of microkernel-based OSes makes it easier to build secure-by-design system software, which is of critical importance for connected devices that have the potential to cause physical damage to infrastructure or even bodily harm to people. In conclusion, systems following the microkernel-based approach are much more suitable for security-critical use cases such as IoT.

3.2.2 Secure-by-default hardware architecture

We believe that microkernel-like ideas can also improve hardware design. Hardware engineers often need to incorporate third-party components into their systems, much like software developers use third-party libraries. For example, system-on-chip (SoC) designers regularly integrate intellectual property blocks from various sources, but not all of them may be trustworthy. A prime example is off-the-shelf wireless modems, which include their own processors to run firmware that implements inherently complex communication protocols. Due to this complexity, the modem may be susceptible to over-the-air attacks that can give the attacker full access to the modem's processing capabilities. With such a foothold, the attacker can exploit the input/output (I/O) bus to which the modem is attached in order to access other building blocks. Being able to access the bus from the device side (i.e., originating from the modem) is both common and necessary. The modem, and many other I/O devices, need to be able to directly read and write the main memory of the computer system via Direct Memory Access (DMA). This ability is essential for maximum performance and high efficiency, as it frees the application processor from tending to data-intensive I/O operations. But it can also be a security risk. Successful attacks via a compromised modem on the application processor's private memory regions have been reported, which constitutes a worst-case scenario as it results in full compromise of the OS and all user data stored in memory. It is important to note that, if an attacker successfully exploits a hardware building block with unrestricted bus access to memory outside the device itself, all software-based defense strategies of the OS and applications are rendered useless; they might even result in cryptographic secrets being leaked.

The lack of isolation between hardware components (i.e., intellectual property blocks of an SoC) is the root cause for these types of security issues. As a defense, operating systems such as Linux, Windows, and macOS make use of a hardware feature called I/O Memory Management Unit (IOMMU). It enables the OS to configure for each hardware device the memory regions that can be read or written by the device. Thus, an IOMMU prevents unrestricted memory access of a device such as a modem via the I/O bus. However, many embedded computers and microcontroller platforms that are in principle suitable for integration into an IoT device lack such a hardware security feature.

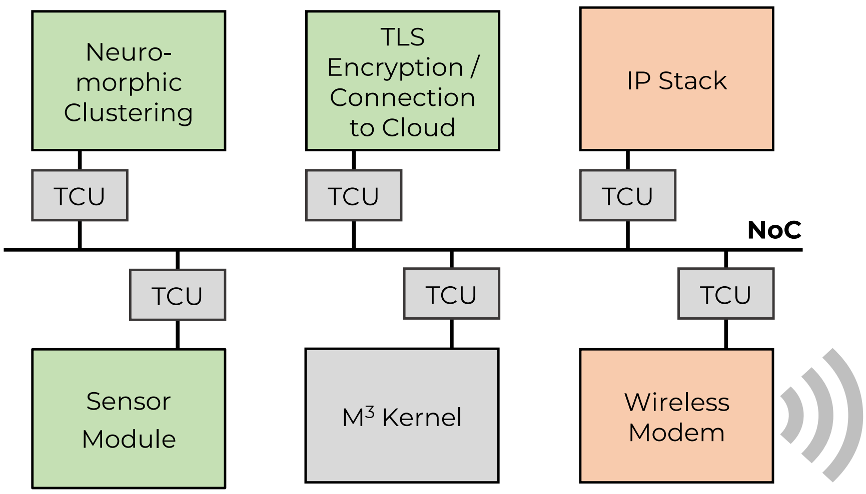
Co-funded by the Horizon 2020 Framework Programme of the European Union

Another problem with IOMMUs is that they need to be properly managed by the OS and that they represent an additional security mechanism that is only applicable to I/O devices. Thus, it forces the developers of the OS and the device drivers to add more complexity to the system software than what is already needed for isolation of OS components (in a microkernel-based architecture) and the applications (in any architecture except for the simplest of RTOS). A unified approach to management and enforcement of all isolation would reduce the overall complexity, as it is the same for all active components in the system.

3.2.3 M³ hardware/software co-design

To advance the state-of-the-art in computer-architecture security and to meet the efficiency requirements of IoT use cases, the isolation-by-default approach of microkernel-based systems should be applied to the hardware structure of the computer system, too. Similarly to how a microkernel isolates mutually distrusting software components, a small and trusted hardware component placed between the device and the I/O bus could limit the damage that a faulty or compromised hardware building block can do to the rest of the system. If carefully designed and integrated with the network on chip (NoC), this hardware component may also allow for simpler integration of third-party intellectual property blocks.

Barkhausen Institute (BI) contributes to iNGENIOUS a microkernel-based OS called M³ [53] and a computer architecture that has been co-designed together with the OS. This architecture is suitable for highly integrated SoC platforms based on *tiles* connected via a NoC. As shown in Figure 3.6, the tiles do not access the NoC directly, but through a small hardware



component called Trusted Communication Unit (TCU). There are two types of tiles:

- **Processor tiles** contain a general-purpose processor that can run software.
- **Accelerator tiles** contain a special-purpose compute unit that may be a fixed-function data processor which is not programmable at all, or a

Co-funded by the Horizon 2020
Framework Programme of the European Union

digital signal processor (DSP); or they might be what is considered an I/O device in classical computer architecture.

Figure 3.6. Microkernel-based approach applied to hardware in the M^3 platform

The microkernel runs on one dedicated processor tile, while other components of the OS and application software are assigned their own, separate processor tiles. Hence, this dedicated tile is called the *Kernel Tile*, whereas all other tiles are *User Tiles*. Accelerator tiles are user tiles as well.

The TCU that sits between each tile and the NoC is a data movement engine that enables message passing between tiles. It also allows the compute unit of the tile to directly access memory that is attached to another processing tile or a global Dynamic Random Access Memory (DRAM) tile. All these data transfers must be done via the TCU, as no other communication links exist in the hardware. The TCU will only allow communication or memory transfers between two tiles, if an *endpoint* specifying the target tile has been configured within the TCU. Endpoints allow either message exchange or reading and/or writing a specific region of remote memory. The endpoint configuration of a TCU cannot be altered by the tile itself, but only the microkernel running on its dedicated kernel tile. The kernel tile is privileged in the sense that it is the only tile from which TCUs can be (re-)configured via the NoC. Thus, the microkernel running on the kernel tile is the only component in the system that can manage communication channels and endpoints for remote memory access. Each TCU participates in enforcing the system-wide access-control policy, which is expressed as either existence or non-existence of endpoints that allow respective message and memory transfers between tiles. This enforcement is done in hardware.

Outsourcing access-control enforcement to the TCUs has two advantages. First, access rights of software running on user tiles with general-purpose processors can be policed without involving the M^3 microkernel, thereby increasing efficiency of communication control. Second, any other kind of user tiles with a hardware accelerator or I/O device can be connected to the NoC and managed in the same way. Accelerators and I/O devices become devices as "1st-class citizens" that are subject to a unified management and access control regime supporting both software and hardware components.

3.3 Secure Cooperation of IoT Devices and the Cloud

A secure-by-default computer and OS architecture can greatly increase the overall security of individual computers, but it is only one building block. In a distributed system like the IoT, each individual device must be secure and reliable. IoT devices do not just “communicate” with each other, but they *cooperate* towards a common goal. For example, an IoT device with a sensor takes measurements of the physical world (e.g., water level in a tank); the measurement data is sent over a communication channel to a control computer, e.g., in an MEC or remote data centre; then, the control computer processes the sensor readings and may direct another IoT device with an actuator to perform some action in response to the sensor reading (e.g., increase or decrease the flow of water into the tank). The sensor, actuator, and control computer are generally distributed *nodes* connected via an IoT network. It is this distributed nature of IoT systems that makes them powerful, but also susceptible to failures and malicious attacks, because the nodes and the communication links between them represent a more complex system with a larger attack surface.

Cooperation between IoT nodes is only possible, if these nodes can rely on each other. For example, the control component needs to be sure that the measurements received from the sensor are correct, whereas the actuator must trust the control computer to only send valid commands. The control computer also needs to be sure that it is talking to the right actuator and that it performs the action as commanded. It follows that nodes in the IoT network need to trust each other for the overall system to function correctly and this trust can be transitive.

A node can be trustworthy, if it behaves correctly. Correct behaviour can be assumed (and therefore be trusted in), if the node is correctly identified and shown to be in a correct state (i.e., not a manipulated replacement, hacked, failed self-test, etc.). Since sensors and even the simplest actuators in an IoT system require computation and communication capabilities, they typically run software that determines their behaviour. It is therefore not only necessary to identify the (type of) device to know it will behave correctly, but the identity and integrity of the on-device software (and potentially data) must be known, too. Moreover, IoT devices require occasional software updates in order to enable new functionality or to fix previously unknown vulnerabilities. The same applies to software stacks running in a MEC or remote data centre. Thus, to remain trustworthy, IoT nodes must be identifiable and able to replace their software, but also secure from manipulation.

A hardware Root of Trust (RoT), in conjunction with a Trusted Execution Environment (TEE), enables identification of software and verification that the integrity of the software has been preserved. Remote Attestation (RA) enables one node to assess the trustworthiness of another node in the distributed system. These *Trusted Computing* concepts are explained in the following.

3.3.1 Root of trust

A Root of Trust (RoT) is some part of a computer system that an attacker cannot change. Typically, it must be part of the hardware and integrated into it in a tamper-proof way, because most threat models assume that an attacker can execute arbitrary code or even has some kind of physical access. A key function of the RoT is to be able to identify the system it is part of. For example, an RoT can be the private part of a signature key pair. This private key is embedded into the hardware in such a way that it can be used to create signatures, but not read by any software. If it were possible to extract the private key, the device identity could be faked and this must be prevented.

3.3.2 Trusted execution environment

Co-funded by the Horizon 2020 Framework Programme of the European Union

A Trusted Execution Environment (TEE) allows execution of a program and keeps this program and its state secure from whatever is happening outside the TEE. To enable the program in the TEE to interact with the outside world, communication channels are used to provide a controlled way for input and output. Some systems also use the term *enclave* to describe a similar idea, where the TEE is embedded into the address space of an application process. The non-enclave parts of this process (and the OS) are therefore outside the TEE and cannot access the memory within the enclave. In the context of M^3-based platform design, any processing tile can be considered as a TEE.

3.3.3 Remote attestation

Remote Attestation is a cryptographic protocol that is executed between two computer systems: a *Challenger* (or *Verifier*) and an *Attester* (or *Prover*). The challenger wants to learn about the state of the attester, which can provide trustworthy state information by leveraging its RoT to prove to the challenger that certain software is running in a TEE on that system. Since the signature key of the RoT only exists in one device (the attester), this device can be identified. The challenger can derive from the identity the properties of the attester device (based on how it is constructed) and can therefore establish trust in the security of the TEE. The signature includes a *Measurement* of the program running in the TEE. For example, this measurement can be a cryptographic hash of the program's binary code and, if needed, relevant configuration. Based on the knowledge about the construction of the attester device, its TEE, and what program is running in the TEE, the challenger can decide to trust the attester or not.

3.4iNGENIOUS Innovation

3.4.1 Neuromorphic context-based data clustering at edge

The basis of Neuromorphic Networks was described in Section 3.1. The iNGENIOUS Transport UC uses neuromorphic networks for the collection of none-biased, balanced data distributions. As written in Section 3.1: “Neuromorphic networks are massively parallel”. This makes them extremely attractive for streaming data analysis. In the use case, iNGENIOUS utilizes neuromorphic networks for data statistics on known clusters and novelty detection. This approach revolutionizes data engineering which is the basis of data science. Typically, data scientists have either too little or too much data, but seldomly unbiased balanced data. The transportation use case leverages neuromorphic cluster networks for long-term studies to measure and collect context-specific cluster space of a given application. Figure 3.7 illustrates this concept:

Co-funded by the Horizon 2020 Framework Programme of the European Union

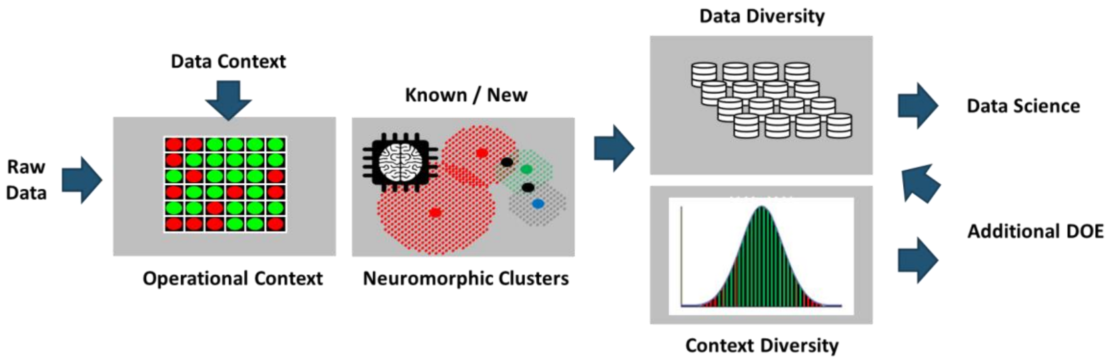


Figure 3.7: Context Based Data Clusters

Data clusters are statistic signal features with similar patterns. These data clusters can be associated with states, which are referred to as context. A state or context can be an operation mode (On/Off/Standby) or another Sensor Signal Range (Small/Mid/Large), or multiple signal ranges (Speed, Acceleration, Temperature, etc.). Any context which has an influence (correlation) on a pre-filtered signal is worthwhile considering in data diversity analysis.

Context information can also be considered as data labels, which will be needed for effective machine learning implementations.

Today, mode data engineering is done by Design-of-Experiments. The data scientists anticipate which contexts affect target signals, and the data scientist introduces artificial process variations and faults and records the outcome. Typically, design of experiments is never comprehensive, or so data intensive that the massive amounts of data squash the data scientist. Old fashioned data engineering typically consumes 80-95% of the engineering resources for less than average quality datasets.

The iNGENIOUS Transport UC is a game changer in how data is collected, mapped, labelled, and recorded. The context-based neuromorphic data logger is an edge server which samples multiple sensor inputs for context generation (labels) and signal cluster collection (non-biased distributed data). The data scientist defines the context data and context resolution (small, mid, large, etc.). The data logger will be tasked to record, for example, 100 recordings on similar signals per cluster. A cluster is a signal with targeted statistical similarity. The clusters are auto generated whenever a new signal does not fit into a previously triggered cluster. The recording will be done on multiple devices for periods of time that are application specific. The recorder will analyse how often unique clusters are triggered within a given context. The outcome is a map of contexts which have been triggered over “lifetime”, and a collection of distributed data clusters. In case the context coverage is deemed too small, targeted design-of-experiments can be added to fill these gaps. The whole process of data labelling and unnecessary data filtering is automated in real time. The data scientist can start his work as soon as the data comes in.

It does, however, require some preparation and knowledge on the target application. In iNGENIOUS, we focus on Rail-Health monitoring using vibroacoustic signals recorded at the bearing housing of axles. Vibroacoustic signals of rail carriages are extremely noise polluted. Clustering of raw signals would not be effective. The raw signal is pre-conditioned (Filters, Short-Time Fourier Transform, Peak-Valley Finder) prior to neuromorphic clustering. Figure 3.8 illustrates context-based neuromorphic data clustering applied to this use case.

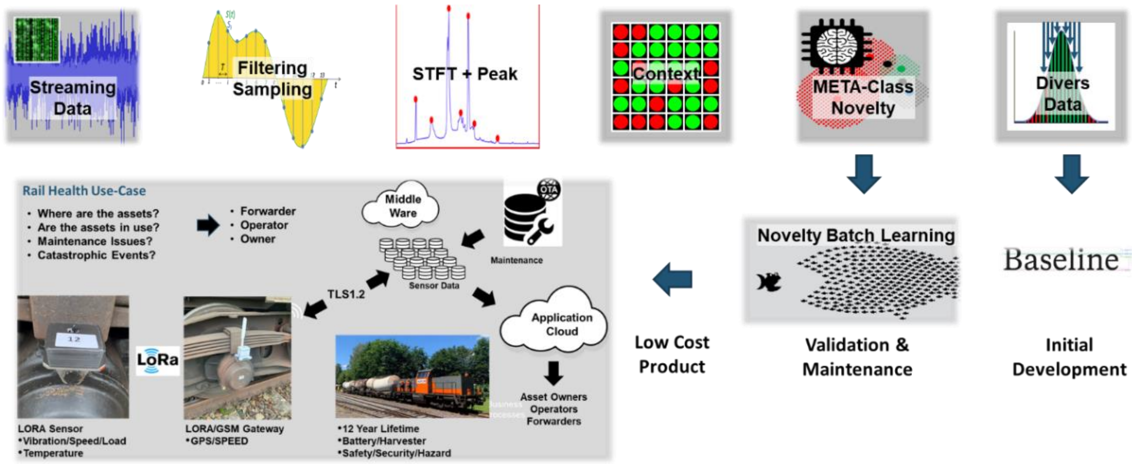


Figure 3.8. Context-based neuromorphic data clustering in the Transport UC

3.4.2 Situational energy optimized edge computing

The situation-based energy optimized edge computing is described in Section 3.1 and the concept will be investigated within Transport UC. An edge sensor that is used for both edge classification and novelty collection is a multi-situational application. Use cases for such sensors are justified during data collection (long-duration testing), or during data validation (0.1% of field sensor used for novelty monitoring and machine learning improvements), or in changing dynamic environments. The sensor must have the capability to identify novelties, their incidence frequency, related raw or feature vectors, and must be able to communicate these higher payloads for cloud-based novelty batch learning.

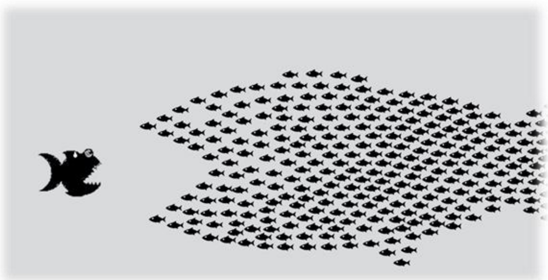


Figure 3.9. Swarm Intelligence via Novelty Batch Learning

The novelty data with available context data is analysed in the cloud, and used if possible, to update the machine learning (ML) application. The revised ML application settings are then distributed to the edge sensors via firmware updates. The energy required to upload novelty data and download firmware updates, is optimized via multi-modal communication channels if available.

Co-funded by the Horizon 2020 Framework Programme of the European Union

3.4.3 Secure-by-default IoT devices

The use of microkernel-based decomposition of the OS and the application of the underlying design principles to the hardware encourages an isolation-by-default approach to constructing IoT devices. As a result of the TCU-enforced access control, the overall architecture becomes secure-by-design for all kinds of activities at the user tiles, assuming that communication and memory endpoints in the TCUs are configured strictly on a need-to-cooperate basis. For example, if an attacker managed to exploit a network-exposed vulnerability in the highly complex baseband firmware of a modem, the OS and application software running on the general-purpose processor tiles can no longer be compromised, because arbitrary memory accesses (e.g., to the dynamic random access memory, DRAM, tile) is not possible. The TCU restricts what tiles and memory regions the modem can access, whereas the system might be fully compromised in an architecture where devices are directly connected to a shared bus without access control mechanisms.

3.4.4 RoT and RA implementation for M^3-based platform

Barkhausen Institute (BI) will research and develop a minimal Root of Trust (RoT) for their M^3 hardware/software co-designed compute platform and integrate it into the FPGA prototype. This RoT will serve as the trust anchor for Remote Attestation (RA) of software components running inside a TEE (i.e., user tile with general-purpose processor) of an IoT device based on the M^3 platform. The design goal for this RA implementation is that it can be validated from a node that uses an RoT/TEE implementation found in a MEC or cloud server (e.g., based on a Trusted Platform Module (TPM)). The RA implementation will be integrated with Transport Layer Security (TLS). In combination, the M^3-based TEE with the FPGA-implemented RoT enables verifiable trust between an IoT device and a cloud endpoint, both of which are connected via a TLS-secured communication channel. After establishment of the TLS connection, the cooperating programs at both sides of the connection will be guaranteed that the identity and integrity of the respective peer has been successfully validated using RA. The RoT and RA implementation can also be used as a mechanism for ensuring software-update integrity. It can also ensure that no version mismatch between the peers goes undetected, which could impair reliability of the cooperating programs running on the IoT device and the cloud server. The development of these M^3 platform features will be performed with the Transport UC in mind.

Co-funded by the Horizon 2020 Framework Programme of the European Union

3.4.5 Relation to UCs

Similarly to the exercise performed in Section 2.4.4, Table 8 below lists the main platform and local computation functionalities introduced in the previous sections, and for each of them how they map with the various relevant iNGENIOUS UCs in terms of innovation maturity.

Table 8. NG-IoT platform functionalities mapped to UCs

Use Cases → / Functionality ↓	Factory	Transport	Port Entrance	AGV	Ship
Neuromorphic context-based Data Clustering		Innovation Demo			
Situational energy optimized computing		Innovation Concept			
Secure-by-default IoT devices		Innovation Demo			
RoT/RA for M³		Innovation Demo			

Co-funded by the Horizon 2020 Framework Programme of the European Union

4 Immersive Devices and Applications

A new generation of devices and applications will arrive to the market in the new 5G networks. These will provide a more friendly and efficient environment to operators who will be able to manage systems and devices with precision and reliability. One of the most exiting use cases of this new area is the Remote Driving, or Tele-operation Driving (ToD), application.

Within WP3, iNGENIOUS will develop and integrate immersive devices such as *Head-Mounted Displays* and *Haptic Gloves*, which are key enablers for improving operation of IoT systems in industrial and logistic environments such as the one targeted to be addressed by the AGV UC. Two practical applications are foreseen within this task:

- **Mixed Reality (MR)** will be used with last generation of immersive head-mounted devices to improve the successful deployment of IoT advanced networks, aiming at guaranteeing the quality and coverage of the network while optimising the project cost on next generation supply chain scenarios.

- **The haptic gloves and haptic sensors** application will improve the perception, quality and safety of the remote operators managing AGVs in hazardous locations of industrial areas, which is required to guarantee the operation even in very exceptional situations where the autonomous robots or vehicles cannot operate. The integration of physiological assessment features, enabling the possibility of monitoring parameters such as user's fatigue or stress are also contemplated and will be developed within this task.

All these advancements will be corroborated, demonstrated and validated by experimentation within iNGENIOUS, where the reliability and precision of the remote operation of vehicles or infrastructures will be defined by the required KPIs provided by each device and network component involved in the different services.

Generally, there are several flows of information from/to user immersive IoT equipment with different requirements in terms of data rate and latencies, covering a variety of IoT market categories, as described in introduction of Chapter 2. In Figure 4.1, we can see a reference classification of information flows in the ToD application supported in AGV UC.

Co-funded by the Horizon 2020
Framework Programme of the European Union

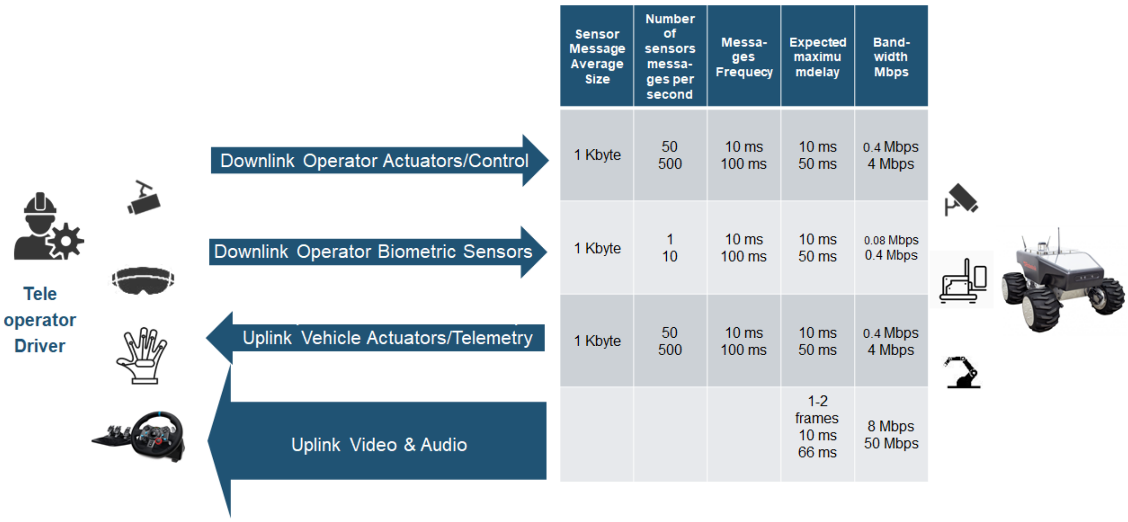


Figure 4.1. Information flows for ToD application service in AGV UC

In this reference classification we can see the following flows:

- **Downlink Operator Actuators/Control**: this information includes the vehicle remote actuators controls for speed, direction, lights, sound, etc.
- **Downlink Operator Biometric Sensors**: this flow of information includes biometric data to verify that the operator is in good shape to continue with the remote operation
- **Uplink Vehicle Actuators/Telemetry**: the information transported in this flow includes the position from Global Positioning System (GPS), and all the in-vehicle sensors information key for the operator like direction, acceleration, speed, radio signal, etc., but also information about actuators in the human feedback
- **Uplink Video & Audio**: this information is low-latency video and audio from the vehicle to the remote operator

In general, these flows above are quite typical and can be used as reference case but other flows may be required in more specific ToD oriented use cases. Depending on more specific requirements on the maximum tolerated delay and the required bandwidth, the vehicle could be operated at higher speed or better precision, so a range is provided for different use case scenarios. These flows could be transported using several QoS characteristics in 5G system, as defined by 5G QoS identifiers (5Qis, or 5GQIs), which are scalars that are used as a reference to 5G QoS. 5GQI value is configured in the Access Node gNB.

In the following of this chapter, we first describe in Section 4.1 the various immersive device solutions that iNGENIOUS will leverage and evolve. Then, in Section 4.2, we discuss the details of the immersive remote indoor cockpit which will be formed by the implementation of the immersive devices and applications in iNGENIOUS. Finally, Section 4.3 presents the innovations and improvements that iNGENIOUS partners currently investigate on that front.

4.1 Immersive Devices

Co-funded by the Horizon 2020 Framework Programme of the European Union

The ToD display system could include several monitors to present the remote video presentation but can also be implemented with 360 degree display devices in order to provide full immersive experiences. These devices should introduce low latency in the video presentation to guarantee the most updated view of the remote environment where the vehicle is moving. The immersive version of the cockpit will be delivered with head-mounted displays that allow the easy implementation of different telemetry compositions to better help the remote operator.

In Figure 4.2 we can see a real picture of the view of a remote ToD operator with the immersive view from a virtual digital twin of the real vehicle. The video presented in the vehicle windows is the low latency video captured by the real cameras installed in the vehicle.

Figure 4.2. Nokia immersive cockpit

The immersive head-mounted devices evaluated in the iNGENIOUS project include two versions, one more low-end using the Oculus Quest 2 and another more professional using the VARJO Extended Reality (XR)-3.

In addition, new devices, such as haptic gloves or other biometric devices, as presented in Figure 4.3, will be incorporated to the ToD cockpits. The gloves, that allow the use of hands gestures, can be used to control remotely the AGVs. They also support some actuators that transmit vibrotactile signals to the operator skin to provide one additional information channel. Other human biometric sensors can provide real time information about the remote driver.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Figure 4.3. Haptic gloves integrated in a complete Immersive Cockpit

The main KPIs for the immersive devices of interest are summarised in Table 9 below.

Table 9. Main KPIs for iNGENIOUS immersive devices

Head-mounted devices	Haptic gloves
• Video resolution • Maximum frame latency • Video encoder type • Video aperture • Video encoding delay	• Sensing frequency • Sensor maximum delay • Sensor accuracy • Actuator maximum delay

In the following, we summarise the main features of these immersive devices considered in iNGENIOUS.

4.1.1 Head-mounted device: Oculus Quest 2

Figure 4.4: Oculus Quest 2 Headset

The main features of the low-end immersive head-mounted device from Facebook, Oculus Quest 2, include:

- Product Dimensions: 191.5 mm x 102 mm x 142.5 mm (strap folded in), 191.5 mm x102 mm x 295.5 mm (strap fully opened up)
- Product Weight: 503g
- Tracking: Supports 6 degrees of freedom head and hand tracking through integrated Oculus Insight technology.
- Storage: 64GB or 256GB
- Display Panel: Fast-switch LCD
- Display Resolution: 1832×1920 per eye
- Display Refresh: 72Hz at launch; 90Hz support to come
- SoC: Qualcomm Snapdragon XR2 Platform
- Audio: Integrated speakers and microphone; compatible with 3.5mm headphones.
- RAM: 6GB
- Battery Life: Between 2-3 hours can be expected based on the kind of content used on Quest 2; closer to 2 hours if playing games and closer to 3 hours if watching media. At any point, the battery status of the headset can be checked in the Oculus App settings or in VR via Oculus Home.
- Charge Time: With the provided USB-C power adapter, Quest 2 will charge to a full battery in about 2.5 hours.
- Interpupillary distance (IPD): Adjustable IPD with three settings for 58, 63 and 68mm.
- Playspace: Stationary or Roomscale supported. Roomscale requires a minimum of 6.5 feet x 6.5 feet of obstruction-free floor space.

Main features of the associated controllers include:

- Dimensions: 9 x 12cm (per controller, includes tracking ring)
- Weight: 126g (per controller, no battery installed)
- Requires 2AA batteries (included in the box; 1 for each controller)

4.1.2 Head mounted device: Varjo XR-3

The main features of Varjo XR-3 include:

- Photorealistic, true-to-life mixed reality powered by low-latency, 12-megapixel video pass-through
- The industry’s highest resolution (over 70 ppd) and the widest field of view (115°)
- Depth awareness for pixel-perfect real-time occlusion and 3D world reconstruction
- The widest-ever colour gamut matches 99% with the sRGB colour space for the most realistic scenes ever produced
- Integrated Ultraleap hand tracking and integrated 200 Hz eye tracking for natural interactions
- Inside-out tracking, offering flexibility for deployments without the need for base stations
- Total comfort with a 3-point precision fit headband, active cooling, and ultra-wide optical design to reduce eye strain and simulator sickness
- Complete software compatibility. Any software is easily ported into Varjo via Open XR 1.0 or Varjo native software development kit (SDK). Varjo also supports OpenVR content. In addition, we have native

Co-funded by the Horizon 2020 Framework Programme of the European Union

support for Unity, Unreal Engine, and dozens of industrial 3D software, including Autodesk VRED, Lockheed Martin Prepar3d, VBS BlueIG and FlightSafety Vital.
- Included in the box: Varjo XR-3 mixed reality headset, 2 headset adapters, 2 power supply units, 6 power plugs (US, EU, UK, China, Australia, Korea) for each power supply, cleaning cloth, Varjo reference marker. The box does not include the SteamVR base stations or controllers.

4.1.3 Haptic gloves

The main features of the haptic gloves include:

- Lycra (antibacterial and fire-resistant) gloves.
- Ultra-low-power high-performance dual core ARM Cortex M4 STM32WB55 processor performing all the real-time low-layer operation at a frequency of up to 64 MHz. The processor supports all Arm single-precision data-processing instructions and data types and embeds a powerful and ultra-low-power radio, compliant with the Bluetooth Low Energy SIG specification 5.2 and with IEEE 802.15.4-2011.
- Ten customized vibrotactile Y-axis linear resonant actuators (LRA) arranged on the fingertips and the palm of the hand. Actuators have a diameter of 10 mm, a resonant frequency of 205 Hz and vibration amplitude of 1.8 G (i.e. the gravitational constant). This provides realistic touch feedback for the user in virtual scenarios.
- Seven 9-Axis Inertial Measurement Unit (IMU) for hand and finger tracking. Sample rate is > 200 Hz.
- Four conductive textile zones that detect fingers-hand interaction for gesture-programmable command triggering.
- Built-in 600 mA Li-Po battery. Under normal work conditions it ensures 5 hours of use.
- One micro Universal Serial Bus (USB) port for battery charge and data communications.
- Wireless communications through Bluetooth 5.0 Low Energy with customized firmware for plug & play for any mobile device. Sample rate is > 72 Hz.
- Compatible with Oculus, Google, HTC headsets.
- Software development kit (SDK) available for Unity and Unreal platforms.

4.2 Immersive remote indoor cockpit

The implementation of the immersive devices and applications in iNGENIOUS will form the immersive remote indoor cockpit. More specifically, the ToD application will run in the MEC, with the purpose of redirecting User Datagram Protocol (UDP) messages associated to the Remote Driving Protocol. It requires an additional piece of software installed in a personal computer (PC) in order to enable direct connection to the MEC instance and that will be used as a cockpit in order to be able to remotely drive the vehicle and to get the telemetry data information into the head-mounted device display.

Co-funded by the Horizon 2020 Framework Programme of the European Union

How this MR cockpit works as well as the involved devices, the hardware, software and connectivity requirements and options are described in the following.

4.2.1 Cockpit role in ToD application

Below we describe the steps included in the ToD application for the AGV UC and the role of MR cockpit:

- Automated routes are defined for the transport of goods in the maritime port terminal. AGV follows this close loop programmed route with stopovers for the loading/unloading of assigned bays.
- The AGV will be equipped with 3 cameras of 120° view each. By placing them in a position forming a triangular structure, a 360° view can be formed. In addition, the AGV will also have integrated proximity sensors to monitor its route and automatically detect objects from an immersive remote indoor cockpit wirelessly.
- The AVG is further equipped with either a 5G sub-6GHz modem or a 5G millimetre wave (mmW) modem for uplink and downlink communication to a public 5G sub-6GHz or private mmW gNB, respectively.
- These public gNBs send the data communication to the MEC and from there to a 5G core network (SA/NSA), to process the data. This data is received in the cockpit designed for remote driving. From the cockpit data can be sent to the AGV, in the opposite direction. Main objective of the use case is that self-driving AGV loads/unloads goods from origin to destination and the immersive remote indoor cockpit allows the operator to take full control of the AGV remotely and perform alternative missions in non-automated zones.
- During the round-trip the AGV sends real-time positioning and status updates to the network infrastructure that are acknowledged by the immersive remote indoor cockpit.
- The immersive remote indoor cockpit provides telepresence and controls the AGV wirelessly, by an operator fully equipped with forehead Augmented Reality (AR) glasses (for the 3D view & dashboard of the real scene), haptic gloves and sensor strap/trackbands (i.e., accessories that can be added to the glove, embedding an IMU sensor for movement tracking), along with a steering wheel and pedals.
- Through the haptic gloves and sensor strap, tactile sensations are felt by the operator during the mission and in the event of immediate risks. The immersive remote indoor cockpit is additionally in charge for the translation of these sensations to a far-edge MEC via fixed fibre (IP access) and/or 5G Hotspot (FastMile 5G Gateway: sub-6GHz & mmW modems) wirelessly. Haptic gloves and sensor trackbands will also capture hand-arm displacement during the remote driving and register biometric signals that provide information about driver's psychological and physical status.

4.2.2 Cockpit hardware requirements and components

Co-funded by the Horizon 2020 Framework Programme of the European Union

The following equipment is required to setup the cockpit infrastructure into the remote driver premises:

- Microsoft Windows 10 based PC
  - Intel i7 or better
  - NVIDIA GTX1070 GPU (or better) graphic card
  - 32 GB RAM
  - 500 GB, or better SSD DISK
  - 4K Display 32", or better
- Oculus Quest 2 or Varjo XR-3 headsets and haptic gloves (see Section 4.1)
- Logitech G29 Steering Wheel and Pedals [54]

Figure 4.5: Logitech G29 Steering Wheel and Pedals

- Logitech Driving Force Shifter [55]

Figure 4.6: Logitech Driving Force Shifter

4.2.3 Cockpit software requirements

- Recent NodeJS installation (v11 or newer) is recommended.
- Oculus driver version: To set up the Oculus Rift S or Rift, the Oculus app needs to be downloaded to the target PC.
- NVIDIA drivers are also required.

Co-funded by the Horizon 2020 Framework Programme of the European Union

4.2.4 Cockpit connectivity requirements

- Logitech Shifter should be connected to the wheel using the cable which is devoted to such purpose.
- The wheel has to be connected to an external power supply and should be configured with the switch on its top part set to PS3 so as the events can be correctly processed by the PC

4.3 iNGENIOUS Innovation

4.3.1 MR cockpit

Using remote driving equipment, Nokia will recreate a vehicle cockpit to be used as a tool for driving vehicles remotely in specific situations, such as AGV UC.

The typical situation in which the remote driving use case make sense takes place when the autonomous vehicle gets stuck and the help of a remote human operator is needed for getting the AGV out of this situation. The cockpit will execute a Unity VR application that will mix some synthetic rendered scenario with real time video coming from a dedicated board in the vehicle which is streaming UDP H.264 video. Another application, based on NodeJS (a JavaScript runtime environment) development, will be also running in the cockpit for handling the remote-control protocol in order to send remote driving commands and receive telemetry from the vehicle. This telemetry information is forwarded to the Unity application for rendering in different visual layer information related to speed, steering, position, etc.

The main innovation is based on an integrated end to end protocol for remote driving, integrating immersive video and control commands over 5G edge infrastructure. All the data flows are optimized and protected using a dedicated MEC architecture allowing low latency and agnostic IP addressing schemas by making use of Internal 5G VPN orchestrated at a virtualized Openstack deployment layer located in a server which is extremely close to the radio infrastructure.

Apart from Nokia's cockpit, an experimental cockpit will be used for parallel trials. This experimental cockpit will be based on the Fivecomm 5G cockpit used for teleoperation of 5G connected vehicles. This experimental cockpit is already working with several peripherals. The haptic gloves will be integrated in that cockpit in a first testing phase. For the final version Nokia's cockpit descripted above will be employed.

4.3.2 Immersive haptic gloves

Haptic gloves will be improved to include biometric sensors:

- Blood volume pulse (BVP) will be detected and used as an objective estimation of operator’s heart rate through an optical sensor integrated in the gloves.

Co-funded by the Horizon 2020
Framework Programme of the European Union

- Respiration frequency and amplitude of breathing: A sensor band around the chest will monitor the expansion and contraction of the operator's rib cage.
- Upper body posture monitoring: the IMU sensors embed in the glove, the trackbands (forearm, arm and chest), and the headset will provide information about operator's body position.

Body posture, hand-arm displacement and physiological data will be processed in order to estimate physical and psychological status and suitability of the operator for task performance (i.e. AGV control). In particular, fatigue, drowsiness and stress levels can be detected and a warning can be delivered in case of risk for the operator.

4.3.3 Relation to UCs

Similarly to the exercise performed in Sections 2.4.4 and 3.4.5, Table 10 below lists the main immersive devices and applications functionalities introduced in the previous sections, and for each of them how they map with the various relevant iNGENIOUS UCs in terms of innovation maturity.

Table 10: NG-IoT immersive devices and applications functionalities mapped to UCs

Use Cases → / Functionality ↓	Factory	Transport	Port Entrance	AGV	Ship
Immersive haptic gloves				Innovation Demo	
MR cockpit				Innovation Demo	

Co-funded by the Horizon 2020 Framework Programme of the European Union

5 Conclusion

This document has described the benchmarking and planned innovations of iNGENIOUS with respect to IoT devices and communication beyond 5G. In particular, this deliverable has detailed how the planned innovation related to WP3 will be employed in the use cases (UCs). The document has been divided in three parts, namely, IoT connectivity, ultra-safe low-power dedicated platforms, and immersive devices and applications.

With respect to the connectivity, this document has surveyed the state-of-the-art (SoA) of several air interface technologies and respective communication devices in the framework of the WP3 task to explore IoT connectivity towards contribution to the Next-Generation IoT (NG-IoT) connectivity design. In particular, cellular (LTE-based, such as LTE-M/NB-IoT, and NR-based) and non-cellular (LoRa, Sigfox, etc.) solutions will be considered to leverage in iNGENIOUS UCs. At the same time, the ongoing 3GPP work in NR RedCap, NTN and NR IIoT was examined in detail as it will be of great focus within WP3 planned technical research for concept innovation opportunities. Moreover, the innovations of iNGENIOUS related to connectivity include the design of low-latency flexible PHY/MAC architecture to operate on a tile-based hardware platform, as well as the development, integration and validation of a versatile 5G communication modem solution. Both of these innovations are planned to be demonstrated in the UCs.

Regarding the ultra-safe low-power dedicated platforms, this document covered first the challenge of optimizing power consumption to maximize the time a device can remain active in the field and the functionality provided by neuromorphic networks for highly optimized parallel, real-time and low power computing with reduced transistor count. The innovations of iNGENIOUS, related to this functionality, include the concept of situational energy optimised computing and the neuromorphic context-based data clustering at edge. The latter innovation includes a context-based neuromorphic data logger, sampling multiple sensor inputs for context generation and signal cluster collection, and is planned to be demonstrated in the UCs. In addition, the document covered the challenge to harden the device against malicious attacks to keep the IoT as a whole trustworthy and the aspect of careful design of both hardware and software. Isolation-by-default architectures in combination with a hardware root-of-trust which can significantly increase the level of security, reliability, and trustworthiness are the respective innovations of iNGENIOUS which are also planned to be demonstrated in the UCs.

In relation to immersive devices and applications, the document has covered the tele-operation driving use case and the related flows of information from/to user immersive IoT equipment which is expected to evolve operation of IoT systems in industrial and logistic environments. The different immersive device solutions that iNGENIOUS will leverage and evolve were described, including head-mounted displays and haptic gloves. Moreover, the document covered the details of the immersive remote indoor cockpit which will be formed by the implementation of the immersive devices and applications. The innovations of iNGENIOUS related to these immersive services include the design of two practical applications, namely, i) mixed reality (MR) application using immersive head-mounted devices to guarantee the quality and coverage of the network while optimising the supply chain cost, and ii) haptic gloves and haptic sensors application to improve the perception, quality and safety of the remote operators managing AGVs in hazardous locations of industrial areas. Both these innovations are planned to be demonstrated in the UCs

Co-funded by the Horizon 2020
Framework Programme of the European Union

References

[1] iNGENIOUS Deliverable 2.1 Use Cases, KPIs and Requirements. [Online]. Available at: D2.1
[2] https://www.digi.com/resources/documentation/Digidocs/90002002/Concepts/c_device_types.htm?TocPath=Zigbee%20networks%7CZigbee%20networking%20concepts%7C_____1
[3] iNGENIOUS Deliverable 4.1 Multi-technologies network for IoT.
[4] https://www.etsi.org/newsroom/press-releases/1839-2020-10-etsi-launches-dect-2020-new-radio-interface-for-iot
[5] https://zigbeealliance.org/solution/zigbee/
[6] https://www.bluetooth.com/learn-about-bluetooth/
[7] https://www.thethingsnetwork.org/
[8] https://www.helium.com/
[9] O. Liberg et al. Cellular Internet of Things – Technologies, Standards, and Performance. 2017.
[10] https://www.telekom.com/en/media/media-information/archive/bye-bye-3g-now-lte-is-coming-for-everyone-608220
[11] https://www.androidcentral.com/end-3g-begins-vodafone-netherlands-shuts-down-its-networks
[12] https://www.commsupdate.com/articles/2020/02/03/3-denmark-to-start-phasing-out-3g-in-2020/
[13] https://www.itu.int/pub/R-REP-M.2410-2017
[14] https://www.itu.int/pub/R-REP-M.2412-2017
[15] https://www.ericsson.com/en/blog/2021/2/how-5g-integrates-tsn-systems
[16] https://www.5gtechnologyworld.com/iot-how-5g-differs-from-lte/
[17] 3GPP SP-160832 "New WID: 5G System Architecture - Phase 1", SA WG2, Meeting: SA#74
[18] 3GPP RP-192313 "Revised WID on Additional enhancements for NB-IoT", Nokia, Meeting: RAN#86
[19] 3GPP RP-191356 "Revised WID for Additional MTC enhancements for LTE", Ericsson, Meeting: RAN#84
[20] 3GPP SP-180614 "Updated SID: Cellular IoT support and evolution for the 5G System", SA WG2, Meeting: SA#80
[21] 3GPP SP-181118 "New WID: Cellular IoT support and evolution for the 5G System", SA WG2, Meeting: SA#82
[22] 3GPP RP-193264 "New WID proposal: Rel-17 enhancements for NB-IoT and LTE-MTC", Huawei, Meeting: RAN#86
[23] Evaluation of LTE-M towards 5G IoT requirements, Sierra Wireless et al, 2018. [Online]. Available at: https://www.gsma.com/iot/resources/evaluation-of-lte-m-towards-5g-iot-requirements/
[24] E. Rastogi et al, Narrowband Internet of Things: A Comprehensive Study, 2020.
[25] 3GPP TR 21.914, Release description; Release 14.
[26] 3GPP TR 21.915, Release description; Release 15.
[27] 3GPP TR 21.916, Release description; Release 16.
[28] 3GPP TR 37.823, Coexistence between LTE-MTC and NR.
[29] 3GPP RP-191584 "Revised WID: Physical layer enhancements for NR ultra-reliable and low latency case (URLLC)", Huawei, Meeting: RAN#84
[30] 3GPP TR 38.825, Study on NR industrial Internet of Things (IoT).

[31] 3GPP RP-200797 “Revised WID: Core part: Support of NR Industrial Internet of Things (IoT)”, Nokia, Meeting: RAN#88-e
[32] 3GPP RP-201310 “Revised WID: Core part: Enhanced Industrial Internet of Things (IoT) and ultra-reliable and low latency communication (URLLC) support for NR”, Nokia, Meeting: RAN#88-e
[33] 5G Core Networks. Power digitalization. Stefan Rommer, Peter Hedman, Magnus Olsson, Lars Frid, Shabnam Sultana and Catherine Mulligan.
[34] 3GPP TS 23.501, System Architecture for 5G System; Stage 2.
[35] 3GPP TR 38.875, Study on support of reduced capability NR devices.
[36] 3GPP RP-210854 "Enhanced Industrial Internet of Things (IoT) and ultra-reliable and low latency communication (URLLC) support for NR" [Online] https://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_91e/Docs/RP-210854.zip
[37] 3GPP SP-200973 "New WID on Support of Enhanced Industrial IIoT" [Online] https://www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGs_90E_Electronic/Docs/SP-200973.zip
[38] NS3 system-level simulator. https://www.nsnam.org/
[39] 5G-LENA module for NS3. https://5g-lena.cttc.es/
[40] 3GPP, “Study on Channel Model for Frequencies from 0.5 to 100 GHz,” 3rd Generation Partnership Project, Technical Report TR 38.901 V16.1.0, Dec. 2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173
[41] T. Jiang et al., "3GPP Standardized 5G Channel Model for IIoT Scenarios: A Survey," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2020.3048992.
[42] R. Bomfin, A. Nimr, M. Chafii and G. Fettweis, "A Robust and Low-Complexity Walsh-Hadamard Modulation for Doubly-Dispersive Channels," in *IEEE Communications Letters*, vol. 25, no. 3, pp. 897-901, March 2021, doi: 10.1109/LCOMM.2020.3034429.
[43] https://www.ni.com/de-de/support/documentation/supplemental/16/labview-communications-lte-application-framework-2-0-and-2-0-1.html
[44] https://www.ni.com/de-de/support/documentation/supplemental/16/labview-communications-802-11-application-framework-2-0-and-2-0-.html
[45] A. Nimr, M. Chafii, and G. P. Fettweis, “Unified Low Complexity Radix-2 Architectures for Time and Frequency-Domain GFDM Modem,” IEEE Circuits and Systems Magazine, vol. 18, no. 4, pp. 18–31, Fourthquarter 2018.
[46] 3GPP RP-130848, “Low cost & enhanced coverage MTC UE for LTE”, Vodafone, Meeting: RAN#60
[47] 3GPP RP-161879, “Requirements for single receiver UE based on category 1 for LTE”, Qualcomm, Meeting: RAN#73
[48] 3GPP TS 38.214, NR; Physical layer procedures for data.
[49] 3GPP RP-200938 “Revised WID: UE Power Saving Enhancements”, MediaTek, Meeting: RAN#88-e
[50] 3GPP TS 22.261, Service requirements for the 5G system.
[51] Dynamic Vision Computing. [Online]. Available: https://www.youtube.com/watch?v=LauQ6LWTkxM

Co-funded by the Horizon 2020 Framework Programme of the European Union

[52] Simon Biggs, Damon Lee, and Gernot Heiser. 2018. The Jury Is In: Monolithic OS Design Is Flawed: Microkernel-based Designs Improve Security. In Proceedings of the 9th Asia-Pacific Workshop on Systems (APSys '18). Association for Computing Machinery, New York, NY, USA, Article 16, 1–7. DOI: https://doi.org/10.1145/3265723.3265733

[53] Nils Asmussen, Marcus Völp, Benedikt Nöthen, Hermann Härtig, and Gerhard Fettweis. 2016. M3: A Hardware/Operating-System Co-Design to Tame Heterogeneous Manycores. In Proceedings of the Twenty-First International Conference on Architectural Support for Programming Languages and Operating Systems(ASPLOS '16). Association for Computing Machinery, New York, NY, USA, 189–203. DOI: https://doi.org/10.1145/2872362.2872371

[54] https://www.logitechg.com/enroeu/products/driving/driving-force-racing-wheel.html

[55] https://www.logitechg.com/enroeu/products/driving/driving-force-shifter.html

[56] Study on Using Satellite Access in 5G, document TR 22.822, Release 16, 3GPP, Jun. 2018.

[57] Study on New Radio (NR) to support non-terrestrial networks, document TR 38.811, Release 15, 3GPP, Sep. 2020. [Online]. Available: https://www.3gpp.org/

[58] Minimum Requirements Related to Technical Performance for IMT2020 Radio Interface(s), document ITU-R M.2410-0, 2017.

[59] F. Rinaldi, H. Maattanen, J. Torsner, S. Pizzi, S. Andreev, A. Iera, Koucheryavy, Y. and Araniti, G. Non-Terrestrial Networks in 5G & Beyond: A Survey. *IEEE Access*, 8, pp.165178-165200, 2020.

[60] RP-202908, “Solutions for NR to support non-terrestrial networks”, 3GPP TSG-RAN Meeting #90, December 2020.

[61] RP-193235, “Study on NB-IoT/eMTC support for non-terrestrial network”, 3GPP TSG-RAN Meeting #86, December 2019.

[62] Study on Narrow-Band Internet of Things (NB-IoT) / enhanced Machine Type Communication (eMTC) support for non-terrestrial networks (NTN), document TR 36.763, Release 17, 3GPP, March 2021 [Online]. Available: https://www.3gpp.org/

[63] Deutsche Telekom IoT: NB-IoT, LoRaWAN, Sigfox: An up-to-date comparison. [Online]. Available: https://iot.telekom.com/resource/blob/data/492968/e396f72b831b0602724ef71056af5045/mobile-iot-network-comparison-nb-iot-lorawan-sigfox.pdf

[64] K. Mekki et al, A comparative study of LPWAN technologies for large-scale IoT deployment, ICT Express, 2017. [Online]. Available: https://doi.org/10.1016/j.icte.2017.12.005

[65] A. Pelaez, LoRaWAN vs NB-IoT: A Comparison Between IoT Trend-Setters, 2020. [Online]. Available: https://ubidots.com/blog/lorawan-vs-nb-iot/

[66] Pothuganti K. and Chitneni A., A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi. Advance in Electronic and Electric Engineering, pp. 655-662, Research India Publications, 2014.

[67] Zanaj, E.; Caso, G.; De Nardis, L.; Mohammadpour, A.; Alay, Ö.; Di Benedetto, M.-G. Energy Efficiency in Short and Wide-Area IoT Technologies—A Survey. Technologies, 2021. [Online]. Available: https://doi.org/10.3390/

Co-funded by the Horizon 2020 Framework Programme of the European Union

[68] technologies9010022R. Kumar et al. Exploring Data Security and Privacy Issues in Internet of Things Based on Five-Layer Architecture. International Journal of Communication Networks and Information Security, Vol. 12(1), pp. 108-121, 2020.
[69] G.S. Karthick and P.B. Pankajavalli. A Review on Human Healthcare Internet of Things: A Technical Perspective, Springer, 2020.
[70] Solutions for NR to Support Non-Terrestrial Networks (NTN), document TR 38.821, Release 16, 3GPP, Jan. 2020. [Online]. Available: https://www.3gpp.org/
[71] 3GPP TR 38.830, Study on NR coverage enhancements (Release 17).
[72] 3GPP TR 38.802, Study on new radio access technology Physical layer aspects (Release 14).
[73] RP-200954, "Revised WID on NR small data transmissions in INACTIVE state", 3GPP TSG-RAN Meeting #88-e, August 2020.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Appendix A: NR-RedCap

Performance impacts derived from UE reduced complexity techniques

UE reduced complexity techniques have impact on the following KPIs: coverage, data rate, latency and reliability, network capacity and spectral efficiency, power consumption and PDCCH blocking rate. In the following, we describe these impacts, individually, per KPI.

Coverage

Potential downlink performance degradation is expected owing to the device complexity reduction, specially, when the number of Rx branches is decreased, hence affecting the coverage.

To compensate for this effect, the study on RedCap UEs addresses the evaluation of coverage recovery based on link budget evaluations. For instance, coverage analysis for FR1 bands can include in the link budget calculation a 3 dB loss from the potential reduced antenna efficiency due to device size limitations. Further, the assumptions in the Rel-17 study item on Coverage Enhancement (CovEnh) [71] are reused (e.g. link budget template, channel specific parameters, antenna array gain, gNB configuration).

The goal of the NR-RedCap study on coverage recovery is to determine the target performance requirement within a deployment scenario for RedCap UEs. The study shows that, for example, PUSCH and/or Msg3 may need coverage recovery up to 3 dB for FR1, unlike for FR2 where coverage recovery for UL channels is not needed.

In addition, the study addresses the coverage recovery features for PUSCH, PDCCH and PDSCH:

- Coverage recovery for PUSCH data includes cross-slot or cross-repetition channel estimation, lower DM-RS density in time domain, enhancements on PUSCH repetition Type A and/or Type B, frequency hopping or BWP switching across a larger system bandwidth.
- Coverage recovery for PDSCH data includes the use of the lower-MCS table, larger aggregation factor for PDSCH reception, cross-slot or cross-repetition channel estimation, increasing the granularity of physical resource block (PRB) bundling, frequency hopping or BWP switching across a larger system bandwidth.
- coverage recovery for broadcast PDCCH includes PDCCH repetition, compact DCI, new AL of 12, 24 or 32, PDCCH transmission via CORESET or search space bundling, PDCCH-less mechanism for SIB1 and/or system information (SI) message.

Data rate

The achievable downlink data rate is affected by three UE complexity reduction features. First, the reduction in the number of Rx branches leads to a significant loss owing to the reduced number of DL MIMO layers that can be supported. This is applied accordingly to the reduction ratio (i.e. by 50%

Co-funded by the Horizon 2020 Framework Programme of the European Union

when the number of antennas is halved and 75% when is divided by 4). Reducing the bandwidth also impacts this parameter as well as relaxing the maximum modulation order (by ~25% from 256 QAM to 64 QAM and ~33% from 64 QAM to 16 QAM).

There is minor impact from HD-FDD operation on instantaneous data rates for UL or DL, but similarly to TDD, HD-FDD reduces user throughput compared to FD-FDD, especially in case of simultaneous DL and UL traffic.

Latency and reliability

Relaxed UE processing time in terms of N1 and N2 has impact on latency. For downlink transmission, relaxed N1 value impacts how fast HARQ-ACK feedback can be sent after the reception of PDSCH. For uplink transmission, relaxed N2 value impacts how fast PUSCH can be scheduled with respect to the UL grant. How significant this impact is depends on use cases and scheduled number of retransmissions.

In general, applying the other techniques slightly impacts the latency and reliability. For instance, HD-FDD introduces longer latency than FD-HDD, especially in case of simultaneous DL and UL traffic, but in any case, the latency and reliability requirements of NR-RedCap use cases can still be fulfilled at least for one direction (DL or UL). The same happens when relaxing the maximum modulation order.

Only in FR2, the UE bandwidth reduction may result in a longer SSB/SIB1 acquisition time for certain configurations for SSB/CORESET multiplexing patterns 2 and 3. To minimize the SSB/SIB1 acquisition time, it may be beneficial to support an FR2 RedCap UE bandwidth of 100 MHz.

Network capacity and spectral efficiency

A loss in network capacity and spectral efficiency is also expected from the reduction in the number of UE Rx branches, the maximum number of MIMO layers, and the maximum mandatory modulation order. In the first case, the magnitude of the loss depends on the proportion of RedCap UEs, the traffic characteristics and the number of Rx branches.

For the evaluation of such impact, the 3GPP used system-level simulations (SLS) and the assumptions in [72]. To this end, two traffic models were considered: burst traffic model and, optionally, full buffer traffic model. These results can be found in [35].

Power consumption

Reducing the number of Rx branches reduces the instantaneous power consumption thanks to the use of fewer RF chains and less complex multi-antenna processing. However, DL reception time may be longer for large payloads due to reduced spectral efficiency. This is also applicable owing to the bandwidth reduction.

The lower insertion loss of a HD-FDD UE may also enable a higher power efficiency in the transmit chain. Compared to the reference NR modem, half-duplex operation means some components can work in a reduced power state until required. However, HD-FDD may have a negative impact on UE

Co-funded by the Horizon 2020 Framework Programme of the European Union

average power consumption because the UE will be active for a longer time before returning to a lower power light or deep sleep state.

Further, relaxed UE processing time in terms of N1 and N2 may allow for processing with lower clock frequency and lower voltage, leading to a reduction in the UE power consumption.

PDDCH blocking rate

Reducing the number of Rx branches leads to a degradation in the robustness of the signalling information transmitted over the radio interface. In order to compensate for the performance degradation, higher aggregation levels may be needed.

As for the bandwidth reduction, if CORESET is configured according to the RedCap UE capability and shared by both RedCap and non-RedCap UEs, this may also result in increased PDCCH blocking rate. In that case, the impact of an FR2 RedCap UE bandwidth of 50 MHz would be greater than for 100 MHz.

HD-FDD operation may potentially reduce the available PDCCH monitoring occasions when the UE is transmitting rather than receiving.

Estimated cost reduction derived from UE reduced complexity techniques

Table 11 presents the average estimated cost reduction for a RedCap UE as compared to a legacy NR UE when the complexity reduction techniques are studied individually. It can be noted that the major cost reduction is achieved thanks to the reduced number of UE Rx branches inasmuch as it provides an approximate 40-50% cost reduction. Reducing the bandwidth follows this trend with a 32% cost reduction for 20 MHz in FR1 and 50 MHz in FR2. The other UE complexity reduction features do not drop the cost as the previous two. For instance, up to a 10% cost reduction may be achieved when using the HD-FDD Type B mode and up to 6% when relaxing the downlink modulation order from 264 QAM to 64 QAM in FR1 and from 64 QAM to 16 QAM in FR2.

Table 11. NR-RedCap average estimated cost reduction from individual UE complexity reduction techniques. Source: [35]

Reduced # of UE Rx branches*	Reduced bandwidth	HD-FDD operation	Relaxed max # of MIMO layers	Relaxed max modulation order	Relaxed UE processing time
FR1 FDD (2Rx→1Rx) **~37%** FR1 TDD (4Rx→2Rx) **~40%** (4Rx→1Rx	FR1 FDD (100→20 MHz) **~32%** FR1 TDD (100→20 MHz) **~33%** FR2 TDD (200→100	HD-FDD Type A **~7%** HD-FDD Type B **~10%**	FR1 FDD (2→1 layer) **~12%** FR1 TDD (4→2 layers) **~17%** (4→1	FR1 DL (256→64 QAM) **~6%** FR1 UL (64→16 QAM) **~2%** FR2 DL (64→16	In terms of N_1 and N_2: FR1 and FR2 **~6%** In terms of CSI computation time**: FR1 FDD **~5%**

Co-funded by the Horizon 2020 Framework Programme of the European Union

) **~60%** FR2 TDD (2Rx→1Rx) **~40%**	MHz) **~16%** (200→50 MHz) **~32%**		layer) **~11%** FR2 TDD (2→1 layer) **~11%**	QAM) **~6%** FR2 UL (64→ 16 QAM) **~2%**	FR1 TDD **~4.5%** FR2 TDD **~6%**
(*) Assuming the corresponding reduction of the number of MIMO layers. (**) The cost reduction gain is estimated without combination with relaxation in terms of N_1 and N_2.					

However, higher cost reductions can be achieved by combining more than one of these UE complexity reduction techniques. Table 12 presents the average estimated cost reduction resulting from a set of combined features for FR1 and FR2. In general, better results are achieved for FR1 TDD for a given combination as compared to FR1 FDD and FR2 TDD. For instance, reducing the bandwidth to 20 MHz, using 1 Rx antenna, relaxing the maximum modulation order down to 64 QAM for DL and 16 QAM for UL and relaxing the UE processing time leads to a cost reduction of 72% in FR1 TDD, 61% in FR1 FDD and 53% in FR2.

Table 12. NR-RedCap average estimated cost reduction from combined UE complexity reduction techniques. Source: [35]

Combined UE complexity reduction techniques	FR1 FDD	FR1 TDD	FR2 TDD
20/100 MHz*, 1 layer, 1 Rx	57.5%	68.6%	47.5%
20 MHz, 1 layer, 1 Rx, HD-FDD type A	63.4%		
20/100 MHz, 1 layer, 1 Rx, DL 64 QAM, UL 16 QAM	59.8%	70.7%	50.7%
20/100 MHz, 1 layer, 1 Rx, double N_1 and N_2	59.3%	70.0%	50.0%
20/100 MHz, 1 layer, 1 Rx, DL 64 QAM, UL 16 QAM, double N_1 and N_2	61.1%	71.9%	52.8%
20 MHz, 1 layer, 1 Rx, DL 64 QAM, UL 16 QAM, HD-FDD type A, double N_1 and N_2	67.1%		
20/100 MHz, 2 layers, 2 Rx		56.6%	
20 MHz, 2 layers, 2 Rx, HD-FDD type A	39.9%		
20/100 MHz, 2 layers, 2 Rx, double N_1 and N_2	35.4%	58.3%	19.1%
20/100 MHz, 2 layers, 2 Rx, DL 64 QAM, UL 16 QAM		59.6%	20.5%
20/100 MHz, 2 layers, 2 Rx, DL 64 QAM, UL 16 QAM, double N1 and N2		61.1%	23.5%

Co-funded by the Horizon 2020 Framework Programme of the European Union

The main contributors to achieve the expected cost reductions for the individual complexity reduction techniques are included in Table 13 for both RF and baseband blocks.

It is assumed that supporting multiple RF bands may affect the RF cost but not the baseband cost significantly. As for the reduction in the number of UE Rx branches and the corresponding reduction in the number of downlink MIMO layers, eleven cost contributors come into play in both RF and baseband, e.g. filtering, receiver processing blocks, low-density parity-check (LDPC) decoding, etc. It can be noted that for an FR1 UE supporting multiple bands, the baseband cost/complexity reduction may be limited by the case with the highest maximum number of MIMO layers among the supported bands. Then, although reducing the bandwidth may solely reduce the complexity in the baseband side, some sourcing companies involved in the study on RedCap UE agree that the power amplifier cost can also be decreased for the proposed 20 MHz bandwidth in FR1. Regarding the HD-FDD operation mode, the main contributor of cost reduction is the duplexer/switch block. For Type A HD-FDD, a high proportion of the cost saving occurs because the duplexer can be replaced with a switch and a low-pass filter. For Type B HD-FDD, uplink and downlink can share one local oscillator; therefore, some additional saving on RF transceiver can be obtained. Finally, relaxing the number of maximum modulation order in uplink will lead to a cost reduction in the power amplifier, transceiver, ADC/DAC and the UL processing blocks, while in downlink the main contributors are the transceiver, ADC/DAC, LDPC decoding, HARQ buffer, and receiver processing block.

Table 13. NR-RedCap main contributor to cost reduction for UE complexity reduction features. Source: [35]

	Reduced # of UE Rx antennas*	**Reduced bandwidth**	**HD-FDD operation**	**Relaxed UE processing time**	**Relaxed max # of MIMO layers**	**Relaxed max modulation order**
RF						
Antenna array	X(**)					
Power amplifier		X(***)				X (UL)
Filters	X					
Transceiver	X					X (UL y DL)
Duplexer			X			

Co-funded by the Horizon 2020 Framework Programme of the European Union

/Switch						
Baseband						
ADC / DAC	X	X				X (UL Y DL)
FFT/IFFT	X	X				
Post-FFT data buffering	X	X				
Receiver processing block	X	X		X	X	X (DL)
LDPC decoding	X	X		X	X	X (DL)
HARQ buffer	X	X			X	X (DL)
DL control processing & decoder				X		
Synchronization / cell search block	X					
UL processing block				X		X (UL)
MIMO specific processing blocks	X				X	

(*) considering the corresponding reduction of the supported maximum number of MIMO layers
(**) only FR2
(***) some 3GPP sourcing companies agree that power amplifier cost can be reduced by reducing UE bandwidth to 20 MHz in FR1.

UE power saving techniques

In addition to the UE complexity reduction techniques, the 3GPP in [35] has also studied the following UE power saving mechanisms: (i) reduced PDCCH monitoring by smaller numbers of blind decodes (BD) and control channel element (CCE) limits, (ii) eDRX for Radio Resource Control (RRC) Inactive and/or Idle mode, and (iii) Radio Resource Management (RRM) relaxation for stationary devices. Further, the enhancements introduced by Rel-17 Power Saving work item should also be applicable to RedCap UEs.

Reduced PDCCH monitoring by smaller numbers of blind decodes and CCE limits

This study includes the evaluation of three potential schemes for PDCCH monitoring reduction:

- Reduced maximum number of Blind Decoding per slot in connected mode with or without reduced DCI size budget. In Rel-15/Rel-16 NR, the number of BDs per slot is configurable up to the limits defined for different SCS configurations (e.g. 36 for 30 kHz SCS). This scheme proposes to reduce such a number. In Rel-15/16, the total number of different DCI sizes configured to monitor is up to 4 with up to 3

Co-funded by the Horizon 2020 Framework Programme of the European Union

different DCI sizes with cell radio network temporary identifier (C-RNTI).

- Extending the PDCCH monitoring gap to X slots (X>1) in connected mode. In Rel-15/16 NR, the range of PDCCH monitoring periodicity is configurable, ranging from a few symbol(s) to 2560 slots subject to UE capability. This scheme aims to increase the minimum separation between two consecutive PDCCH monitoring occasions, spans or slots with configured PDCCH candidates from 1 to X> slots, where X needs to be specified.
- Dynamic adaptation of PDCCH BD parameters in connected mode. This scheme intends to dynamically adapt PDCCH BD parameters (e.g. maximum number of PDCCH candidates per PDCCH monitoring occasion and minimum time separation between two consecutive PDCCH monitoring occasions).

Table 14 presents the analysis of UE power saving through the mean value of the power saving gain when reducing the maximum PDCCH blind decoding by 25% and 50% for both FR1 (figures in black) and FR2 (figures in brown). In general, it is expected that the power saving gain by BD reduction for cross-slot scheduling is less than the one for the same-slot scheduling. Also, the power saving gain by BD reduction for 1 Rx case is expected to be less than that of the 2 Rx case.

Table 14. Mean value of power saving gain when reducing maximum PDCCH BD by 25% and 50%.

Same-slot scheduling		**1 Rx**		**2 Rx**	
		25%	**50%**	**25%**	**50%**
Instant message traffic model		2.81% 4.20%	5.82% 8.60%	3.05% 4.52%	6.59% 8.98%
Heartbeat traffic model	200ms inactivity timer configuration	1.56% 1.72%	3.25% 3.69%	1.65% 2.13%	3.72% 4.14%
	80ms inactivity timer configuration	1.33% 1.28%	2.92% 2.58%	1.49% 1.99%	3.42% 3.88%
VoIP traffic model		2.59% 3.81%	4.74% 7.43%	2.85% 4.27%	5.66% 8.27%
Cross-slot scheduling		**1 Rx**		**2 Rx**	
		25%	**50%**	**25%**	**50%**
Instant message traffic model		2.58% 3.19%	4.26% 6.17%	3.08% 3.43%	5.7% 6.59%
Heartbeat traffic model	200ms inactivity timer configuration	1.66% 1.30%	2.48% 2.60%	1.95% 1.05%	3.51% 2.11%
	80ms inactivity timer configuration	1.60% 1.24%	2.34% 2.48%	1.69% 0.92%	3.21% 1.84%
VoIP traffic model		2.29% 3.27%	3.20% 6.33%	2.28% 3.38%	4.45% 6.52%

As for the performance impacts, blind decoding reduction has no impact on latency unless the PDCCH blocking rate is increased. In that case, the

Co-funded by the Horizon 2020 Framework Programme of the European Union

average latency is expected to be increased. Scheduling flexibility may or may not be impacted by BD reduction depending on multiple factors (e.g. BW, Subcarrier Spacing (SCS), CORESET size, aggregation level (AL) distribution, channel condition, number of ALs per UE, number of UEs that need to be simultaneously scheduled, DCI size budget reduction, etc). The potential impacts on legacy UEs, in terms of PDCCH blocking rate, when coexisting with RedCap UEs in a shared CORESET depend on the scheduling strategy and system parameters. Depending on the network implementation, if legacy UEs are prioritized over RedCap UEs, there is no coexistence impact on the legacy UEs at the cost of increased latency at the RedCap UE side.

Extended DRX for RRC Inactive and/or Idle

In LTE, the UE may be configured with an eDRX cycle. The UE may operate in eDRX only if the UE is configured by NAS and the cell indicates support for eDRX in System Information (note that there is no System Information indication for NB-IoT). In RRC_IDLE, the eDRX cycle has the maximum value of 2621.44 seconds (43.69 minutes). For NB-IoT the maximum value of eDRX cycle is 10485.76 seconds (2.91 hours).

The applicable parts of eDRX mechanisms for LTE, including use of Hyper SFN (H-SFN), Paging Hyperframe (PH) and Paging Time Window (PTW) are expected to be reused for RedCap UEs. For eDRX cycles below and equal to 10.24 seconds, it is recommended that PTW and PH are not used.

For the upper bound, it is recommended that eDRX cycles in RRC_IDLE are extended up to 10485.76 seconds (2.91 hours), unless RAN4 indicates such eDRX value requires UE to perform RRM on serving cell outside PTW.

Extending the eDRX cycle in RRC_INACTIVE beyond 10.24 seconds is beneficial for UE power consumption, hence recommended, given the following reasons. First, to effectively support the use of Rel-17 Small Data Transmission [73] targeting, for instance, use cases with periodic uplink data such as some industrial wireless sensors. Second, to reduce the signaling from the network point of view since there is a need for less RRC signaling. Finally, it has been proved that there is a power saving gain vs eDRX in RRC_IDLE at least for eDRX cycles in the range from 10.24 seconds up to couple of minutes (UE RRC_INACTIVE mode additionally benefits from less signaling). Based on these results, lifetime of several years would not be achievable in some cases (e.g. 1-minute inter-arrival time) if only RRC_IDLE can be used, because of the signaling overhead.

The potential issues with eDRX extension beyond 10.24 seconds for RRC_INACTIVE are: i) impact on core network procedures (e.g. NAS retransmission), ii) potential handling of different eDRX cycles beyond 10.24 seconds and/or PTWs, one for IDLE and one for INACTIVE, and iii) deciding which node decides and configures the eDRX cycle for RRC_INACTIVE.

RRM relaxation for stationary devices

Rel-16 NR RRM relaxation procedures are taken as a baseline to study further enhancements of neighbour cell RRM relaxation for RedCap UEs in RRC_IDLE and RRC_INACTIVE.

Co-funded by the Horizon 2020 Framework Programme of the European Union

The study includes the definition of the possible RRM relaxation triggers and the candidate RRM relaxation methods for stationary UEs. It is recommended that enabling or disabling RRM relaxation should be under network's control.

RRM relaxation has been studied for all the RRC states (RRC_IDLE, RRC_INACTIVE and RRC_CONNECTED) and for both neighbour cell and serving cell measurements. For RRC_CONNECTED, it is recommended that UEs which are fixed or immobile are considered with higher priority compared to UEs which are slightly moving. Irrespective of RRC state, serving cell RRM relaxation for RedCap UEs is not recommended to be specified.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Appendix B: NTN

Channel Modelling

In TR 38.811 [51] the NTN channel models were developed based on the terrestrial 3GPP models. Outdoor conditions only were considered for satellite operations (indoors as well for HAPS), considering several environments: open, rural, suburban, urban, and dense urban. The models were defined to support frequencies from 0.5 to 100 GHz (sub6GHz and Ka bands) and to accommodate UE mobility up to speeds of 1000 km/h for satellite access and 500 km/h for HAPS access.

For terrestrial models we had a multipath phenomenon caused by objects near the user. For NTN, due to the large distances to the satellites, the different signal paths are almost parallel, which translates into an angular spread very close to zero. For this reason, the large-scale parameters such as line-of-sight (LOS) probability or delay spread will depend on the satellite elevation angle and will differ from the terrestrial case.

For the path loss modelling, different components are considered but the main contribution (the basic path loss model) accounts mainly for the signal's free space propagation and adds clutter loss and shadow fading to account for the attenuation of buildings and objects. Their values are defined for different elevations angles and the two frequency bands (S-band and Ka-band). The rest of components of the path loss model are parameters accounting for atmospheric gases and ionospheric and tropospheric scintillation. The channel models must also account for long propagation delays, Doppler shifts and variation rates and Faraday rotation.

Considering the described parameters, two models were developed. First, a generic frequency-selective model based on the terrestrial model with adaptations to account for the satellite geometry. Secondly, a flat-fading two-state model was used for specific situations implying low frequencies, big elevation angles or close to line-of-sight paths among others.

For link-level simulations, reference clustered delay line (CDL) and tapped delay line (TDL) are developed. Two LOS and two NLOS models are constructed to represent different channel profiles, applicable to different environments and different elevation angles.

Radio Interface

Release-16 study on solutions for adapting NR to support NTN concluded in TR 38.821 [66]. The main objective was to identify a minimum set of necessary features to enable NR support for satellite communication networks. The study concluded changes in architecture as well as physical and higher layer protocol changes.

For the physical layer, link and system level simulations in both S-band and Ka-band were conducted showing that handheld UEs as well as UEs with high transmit and receive antenna gains such as Very Small Aperture Terminal (VSAT) can be served by both LEO and GEO satellites in both S-band and Ka-band. The control procedures studied mainly referred to timing relationships, which are much shorter for terrestrial systems. Long

Co-funded by the Horizon 2020 Framework Programme of the European Union

propagation delays require modifications of many aspects in NR involving downlink-uplink (DL-UL) timing interactions. An offset addition in these timing relationships was considered as a solution. Other issues dealt with uplink power control, adaptive modulation and coding (AMC) and delayed CSI feedback, beam management and polarization support and impact of feeder link switch, with no convergence to particular solutions, leaving room for discussion in the specification. Additionally, based on the NR design, DL synchronization via SSB, random access via PRACH and maintenance for UL timing advance were studied. DL synchronization and PRACH design from Release-15 was found to be robust in the case of GEO and LEO with beam specific pre-compensation of common frequency shift. If no pre-compensation is performed, additional complexity at the UE receiver is needed and enhanced PRACH formats and preamble sequences should be supported. Potential optimizations are considered for the normative phase. To ensure performance for timing and frequency synchronization for UL transmissions given the larger cell coverage, long round-trip time (RTT) and high Doppler, enhancements were also considered for the timing advance (TA) and the UL frequency compensation. TA solutions include autonomous acquisition of TA at the UE if we know the UE location and satellite ephemeris, or TA adjustment based on network indications. Frequency offset estimation and pre-compensation was proposed either to be conducted at the UE side by means of DL reference signals, UE location and satellite ephemeris or to be indicated by the network.

The higher layer protocol stack is divided into the User Plane (UP), responsible for data transmission, and the Control Plane (CP), responsible for signalling. For the UP, the main challenge is again the longer propagation delays. MAC, radio link control (RLC), PDCP and service data adaptation protocol (SDAP) were studied. No necessary changes for the SDAP layer were considered in the end for NTN; the considered enhancements for MAC, RLC and PDCP are presented below.

The discussed MAC problems will require enhancements for random access, DRX, scheduling request, and HARQ. Regarding the RACH enhancements, due to differential delay within the same cell in NTN, the preambles sent by different UEs in the same RACH occasion (RO) may reach the network at different times. Possible solutions studied were proper PRACH configuration in the time domain (time interval between consecutive ROs should be larger than twice the maximum delay within the cell) and preamble division such that ROs that are too close in the time domain are assigned different groups of preambles. Solutions such as frequency hopping and 2-step RACH were left as possibilities for further study. After the Random Access Preamble (Msg1), the UE waits for the Random Access Response (RAR) message (Msg2). The interval of wait time is defined by the response window (*ra-ResponseWindow*), which is very short for terrestrial communications and hence should be modified for NTN support. An offset at the start of the *ra-ResponseWindow* was looked at for NTN. The possibility to extend the window was also studied, left as a point to be discussed during the work item phase. After Msg3, the RRC Connection Request, the *ra-ContentionResolutionTimer* starts. A long enough timer to cover the round-trip delay in NTN would require too much power consumption, reason why an offset for the start of it was considered. Coexistence in UEs having GNSS

Co-funded by the Horizon 2020 Framework Programme of the European Union

and non-GNSS capabilities was also studied. This might be an issue considering the possibility of different random access schemes. A solution would be for the network to separate resources and differentiate these based on GNSS capabilities.

The DRX supports UE battery saving by reducing the PDCCH monitoring time. *drx-HARQ-RTT-TimerDL* and *drx-HARQ-RTT-TimerUL* are the minimum duration before a downlink/uplink assignment for HARQ retransmission is expected for the MAC entity. This parameter configurable range is too small for NTN long delays and an offset was proposed as a solution. In the case that HARQ feedback is enabled or disabled only for a certain number of HARQ process IDs, the UE would monitor the PDCCH for retransmission opportunities that would never happen, wasting energy. A possible solution would be not to start the *drx-HARQ-RTT-TimerDL* if HARQ feedback is disabled and to only start the *drx-HARQ-RTT-TimerUL* If HARQ feedback is enabled for the corresponding HARQ process. If HARQ is disabled, however, blind HARQ transmissions are used and the DRX procedures may have some impact. This could be solved by starting the drx-RetransmissionTimer upon network scheduling via PDCCH so that the UE can sleep between the blind transmissions. Lastly, since the network could start sending DCI allocations to the UE with a period that will rarely coincide with the active time of the UE DRX cycle, an extra delay on top of the RTT will be needed to delay transmissions to the first available onDuration period. This could be avoided by allowing the UE to leave its DRX state at the time when the first possible DCI could be received on PDCCH.

The scheduling request (SR) for UL-SCH resources from the gNB requested by the UE cannot be initiated when the prohibit timer (*sr-ProhibitTimer*) is active. This timer will be active for at least 128ms and initiate a SR. However, the *sr-ProhibitTimer* range must be extended to support NTN since the round-trip delay (RTD) in GEO systems is larger.

HARQ at MAC sublayer allows for error correction and repetition. For NTN, the network should be able to configure the enabling/disabling of uplink HARQ feedback for downlink transmissions per UE and per HARQ process. This could be done via RRC signalling. As per uplink transmissions, the enabling/disabling could be configurable on a per UE, per HARQ process and per LCH basis.

The focus on the RLC layer lied on status reporting and sequence numbers. A status report is triggered when there is a failure in the reception of an AMD PDU, indicated by the expiration of the *t-Reassembly*. This timer should cover the maximum time allowed for HARQ transmission, which will probably be larger than the RTD for NTN. Hence, if HARQ is supported by NTN, an extension of the value of *t-Reassembly* would be needed. Regarding the sequence numbers, airplane connectivity was considered the hardest NTN scenario to achieve the required RLC data rate needed. The possible solution would be to either extend the RLC sequence number (SN) length or to reduce the delays it takes to perform an RLC transmission.

Service Data Unit (SDU) discard and sequence numbers were studied for the PDCP layer. The discardTimer used by PDCP to discard the PDCP SDU when it is not successfully delivered can be configured in a range from 10ms to

Co-funded by the Horizon 2020 Framework Programme of the European Union

1500ms. The election of this value should account for the RTD as well as the number of retransmissions on RLC, keeping in mind that its extension will require more memory for the buffer. Modification of this timer will be discussed during the work item phase. For sequence numbers, the PDCP SN length should also be extended, or the retransmission delays reduced to support airplane connectivity.

For the CP, the study focused on mobility management procedures for both idle and connected mode. Satellites provide very large cells which leads to large tracking areas. Moving tracking areas is hard to manage as there is going to be a big trade-off between tracking area updates (TAU) and the paging signalling load. However, they must be dimensioned to minimize TAU as they are more signalling-intensive than paging on the network. In order to avoid frequent TAU, Earth-fixed tracking areas can be used. For idle mode, terrestrial mechanisms are considered as baseline for mobility. Adaptation issues considered were: too frequent SI update, which can be solved by network implementation; cells sweeping the Earth signalling is not a burden due to frequent TAU for LEO satellites; UEs with low transmission power camping on a cell with a high altitude, which can be avoided by the UE if it is able to identify the GEO cell. Ephemeris and UE location information as well as PCI and frequency information can be of help for the UE to perform cell selection and reselection. For connected mode, enhancements for handover were discussed. The high speed of LEO satellites with respect to a fixed position on earth leads to frequent and unavoidable handovers that may result in significant signalling overhead and power consumption. A potentially large number of UEs needing to perform handover at the same time may lead to even more signalling overhead and service continuity challenges. Movement of LEO satellites with moving beams were considered the main challenge to be addressed. Different trigger conditions for handovers were analysed for further specification: measurement-based, location-based, time/timer based, timing advance value based and elevation angle of source and target cells based.

Co-funded by the Horizon 2020 Framework Programme of the European Union