

Partners: NFI, EAB, AAU, AAP, ATO, EBY, ICC, LTU, TNO, NXW, NGE, OPL, TUD, TIM, UBW, UC3, WIN, TUK, SON



HEXA-X-II D3.2 Deliverable

Initial Architectural enablers

Hexa-X-II

hexa-x-ii.eu



Table of Content



- Objectives
- Use cases
- AI enablers for data-driven architecture
- Network modularisation
- Architectural enablers for new access and flexible topologies
- Network beyond communications
- Virtualization and Cloud transformation (IoT-Edge-Cloud continuum enablers)
- References
- Acronyms



Objectives



Deliverable D3.2 objective

- This is the second deliverable D3.2 from WP3
- The main objective of this document is to make an initial description of the different studies in WP3. Based on the different studies, initial descriptions of the enablers are made. The description includes the reason and motivation why this enabler is important for the 6G architecture

WP3 objectives



The overarching objective of WP3 is to develop a 6G architecture framework and innovative enablers for beyond communications and data driven architecture to power new services, modular cloud-native network for improved signalling and new access and flexible topologies for improved reliability.

The WP objectives are:

- **WPO3.1:** Develop and analyse 6G architecture framework and new innovative enablers for the beyond communications and data driven architecture, identify the requirements a data-driven architecture will have on protocols, interfaces, data, and network nodes.
- **WPO3.2:** Define and analyse solutions that combine cloud technology flexibility with distributed processing nodes into self-contained modules with minimum dependency that can be used to extend and scale the network deployments in stepwise manner.
- **WPO3.3:** Develop and analyse new access for flexible topologies and local communications, including different types of multi-connectivity, node roles and node coordination, as well as design control and management solutions for programmable and context-aware transport.



Use cases

Common use cases



WP3 use cases summary

1. From robots to cobots
2. Sustainable development
3. Massive Twinning
4. Immersive telepresence for enhanced interactions

Use case #1: From robots to cobots



• Brief Overview (Hexa-X, D1.3):

- In environments with multiple robots, the machines need to identify other machines, connect with each other, as well as exchange information and intent, in order to fulfil complex tasks efficiently or better cater to the needs and demands of humans in day-to-day interactions [Hexa-X, D1.2]
- Industrial robotics technologies evolve from traditional industrial robots to co-bots to AI-enabled robots with high or full autonomy.
 - Collaborative robot systems sense, perceive, plan and control independently towards a common goal and without explicit instructions from a human.
 - The robots will be equipped with video cameras streaming to a local compute server for real-time processing as well as equipped with advanced sensing and positioning features.
 - Robots will utilize the connected AI capabilities offered by 6G for situation-aware cooperation and collaboration and assistance.
- Humans interacting with machinery or mobile robots indirectly or directly.
 - Robots learning from humans as a result of interaction on a task, e.g., optimizing execution steps or improving error mitigation and prevention steps. Machinery can perform highly individualized on-demand tasks, enabling lot size one production and fully utilizing novel production methods such as additive manufacturing



Interacting and cooperative mobile robots

• Examples:

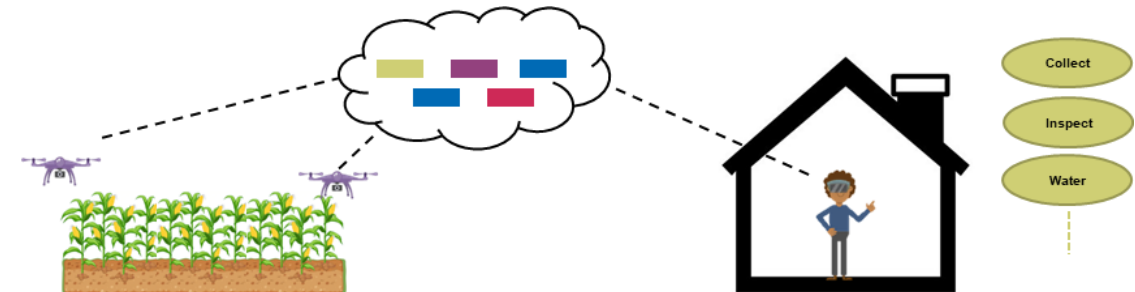
- Interacting and cooperative mobile robots
- Flexible Manufacturing
- Consumer Robots
- Farming use case - farms are usually in remote areas - cobots can provide inspection/caring - virtual presence of farmer can increase the effectiveness (see figure)

• From robots to cobots may be enabled by:

- Distributed Intelligence
- Identification of use case specific APIs to allow cobots run their AI services
- Automated deployment of AI cloud-native functions (with focus on AI agent and AI monitoring), exploring extreme-edge capabilities of cobots to enable distributed AI
- Design of MLOps pipelines for the lifecycle management and continuous monitoring and improvement of AI/ML models that are used by cobots
- Network of networks (e.g., subnetworks, D2D) and trustworthy, flexible topology management due to the mobility of the cobots
- Multi-connectivity for higher resilience
- E2E context awareness management, since the network may adapt dynamically to the cobots' context
- Flexibility in network topologies and resource allocation
- Functional support to meet extreme requirements, i.e., latency, dependability and positioning.
- Close loop control of network functionality

• Requirements :

- High Reliability and Availability
- Low latency
- High Data Rate and High Accuracy
- Mobility



Flexible manufacturing with cobots - farming use case

Use case #2: Sustainable development



- **Brief Overview (Hexa-X, D1.3):**

- The 6G system will be very well suited to provide solutions contributing to meet the UN SDGs or helping verticals to reduce their environmental impact [Hexa-X, D1.2].
- Among other use cases, this use case family addresses the need for inclusion by delivering key digital services, such as health services, to remote or isolated areas. It also focuses on the protection of the environment
- The extreme performance and global service coverage will empower the underserved and bridge the digital divide virtual-realistic remote experiences as well as provide means to monitor and counter-act current and impending environmental challenges.
- Connected intelligence and network of networks will enable operations to be optimized for sustainable performance.
 - Any system addressing the societal and environmental challenges must be trustworthy at its core in order to foster widespread confidence in their operation and execution.

- **Examples:**

- E-health for all
- Institutional coverage
- Earth monitor
- Autonomous supply chains

- **Sustainable development may be enabled by:**

- Privacy-aware data classification collection and learning methods to allow for trustworthy inclusion of sustainable technology development
- Distributed machine learning model training and inference
- Model-driven, privacy-aware, and distributed intelligence solutions to orchestrate many heterogeneous 6G network slice instances
- Network of networks (e.g., NTN and TN integration, subnetworks) for providing extreme global service coverage to remote or isolated areas
- Trustworthy flexible topologies
- E2E context awareness management, since the network may adapt dynamically to the application's context

- **Requirements:**

- Extreme and sustainable performance
- Global service coverage
- Trustworthiness



UN sustainability goals

Use case #3: Massive Twinning



- **Brief Overview (Hexa-X, D1.3):**
 - Massive twinning is designed to lead towards a full digital representation of the environment, extending the use in
 - Production/manufacturing, the management of our environment, in transportation, logistics, entertainment, social interactions, digital health, defence and public safety.
 - A Digital Twin is a virtualized model and real-time representation of an asset in the physical world, i.e., a representation within the digital domain, of the asset's structure, role and Behavior.
 - Some example areas are manufacturing, drastic enhancement of liveability of our cities, the enhancement of the productivity for addressing the "towards zero-hunger" goal, while maximizing the sustainability and the transformation of sectors like health or public security.
- **Examples:**
 - Digital Twins for manufacturing
 - Immersive smart city
 - Digital Twins for sustainable food production
- **Massive Twinning may be enabled by:**
 - Context-aware connectivity, where information about the use case's environment may be gathered at any given time and the system behaviour may be adapted accordingly
 - MLOps pipelines to collect the data for generating and updating the Digital Twin models through continuous monitoring
 - Enhancing the existing infrastructure for data exchange, enabling AI solutions to gather the necessary data for predicting and designing massive twinning models, which will facilitate the efficient flow of data required to train and optimize the AI models
- **Requirements:**
 - High Reliability and Availability
 - High Safety, Maintainability, Integrity
 - Low latency
 - High Data Rate
 - High Location Accuracy



Digital twin used for manufacturing

Use case #4: Immersive telepresence for enhanced interactions



Brief Overview (Hexa-X, D1.2):

- This use case family consists in being present and interacting anytime anywhere, using all senses if so desired [Hexa-X, D1.2].
- As a use case from Hexa-X, immersive telepresence for enhanced interactions covers a large family of use cases including fully merged cyber-physical worlds, mixed reality codesign etc.
- Humans to interact with each other as well as the physical and digital things.
 - Exchanging sensory information through all human senses
 - Extending the capability of the senses
- Enriched experience by the seamless unification of the physical, digital, and human worlds.

Examples:

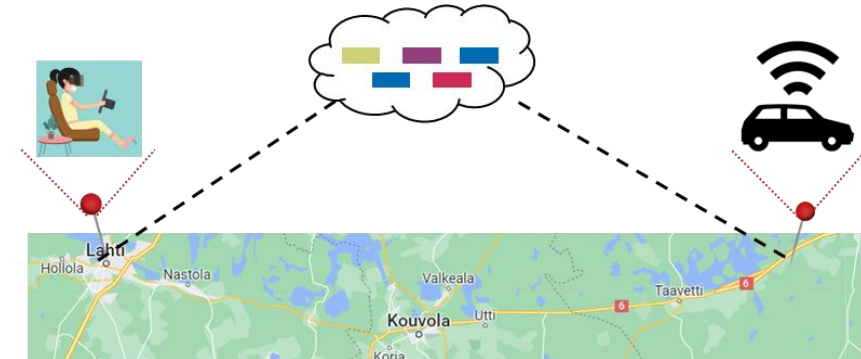
- Smart transportation: autonomous driving, the remote driver has the telepresence of the all the surroundings - increased decision making / working while commuting
- Coworking with the robots and humans - colleagues working together at the digital world whereas in physical world they are just controlling 2 robots
- Immersive education: fully or partially remote education session, where students and instructors will be able to participate in the immersive classroom
- Field robots for hazardous environments/emergencies: In fire management the decisions cannot be automated as they heavily rely on the conditions. Using robots/cobots in emergency response - while providing the telepresence of the fire to the firefighter in a safe place.

Immersive telepresence for enhanced interactions may be enabled by:

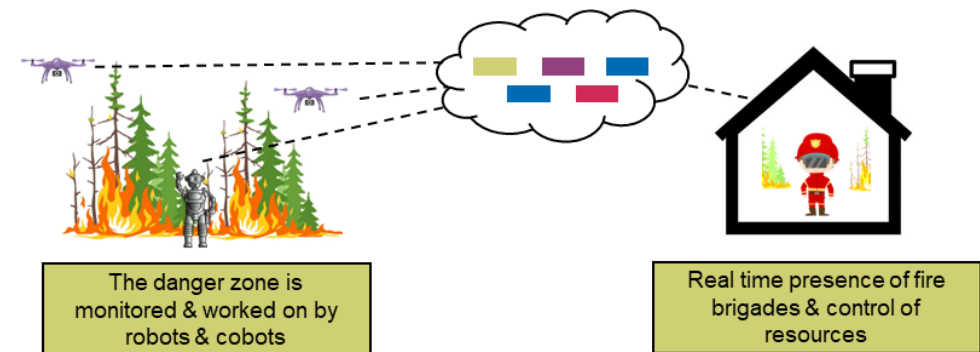
- Subnetworks for providing fast and low-latency communication
- Flexible topologies, especially when user mobility is taken into consideration for seamless mobility and service continuity
- Multi-connectivity for data offloading and higher resilience
- Local information exchange between the nodes of the subnetwork
- E2E context awareness management, since the network may adapt dynamically to the user's sensory context
- Global service coverage
- Runtime scalability to meet performance requirements, e.g., high data rates, low latency, reliability
- Runtime scalability & dynamic function split
- Ultra-high flexibility to dynamically support QoS/QoE
- Exposing the capabilities/data between entities & slices

Requirements:

- High level of flexibility (resources, deployment etc.)
- Lower latency
- Higher bandwidth utilization
- Enhanced reliability and resilience at network level



Smart transportation: the remote driver has the telepresence of the all the surroundings



Field robots for hazardous environments/emergencies:

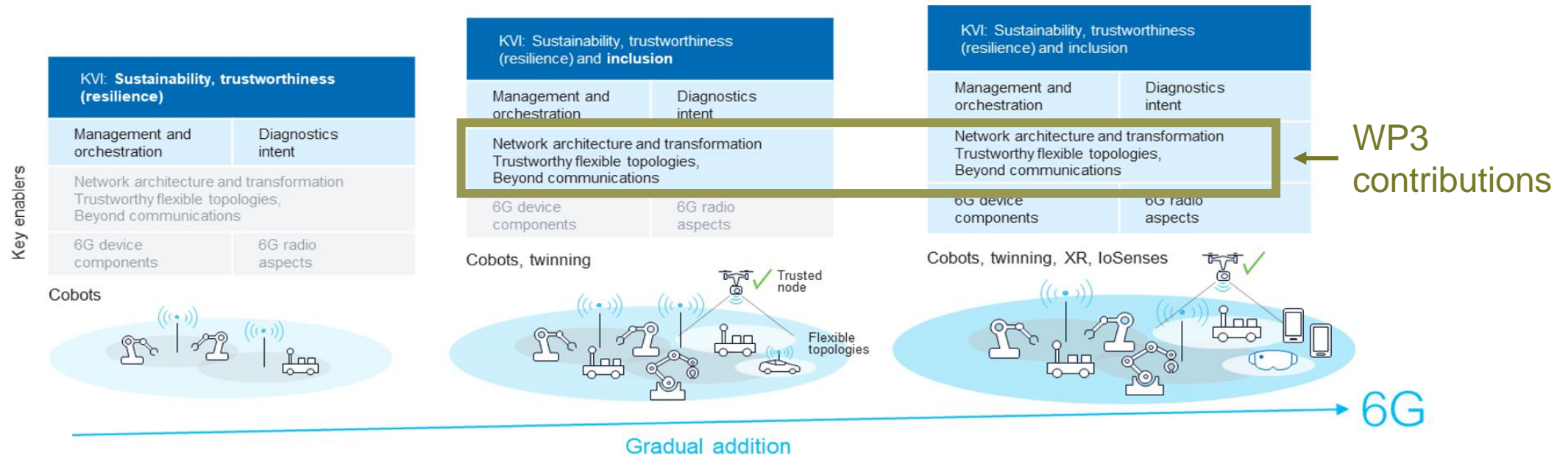


WP3 in Hexa-X-II PoCs

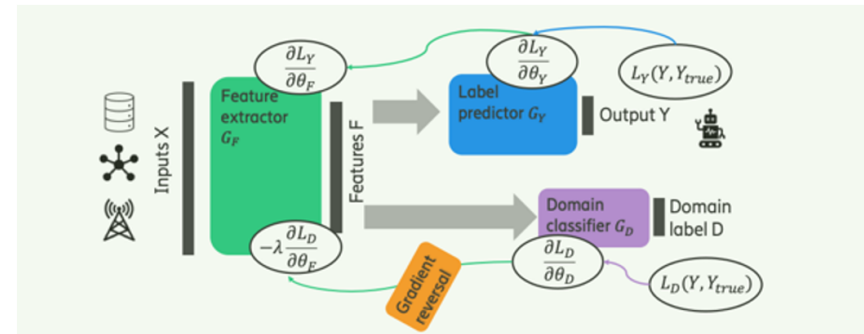
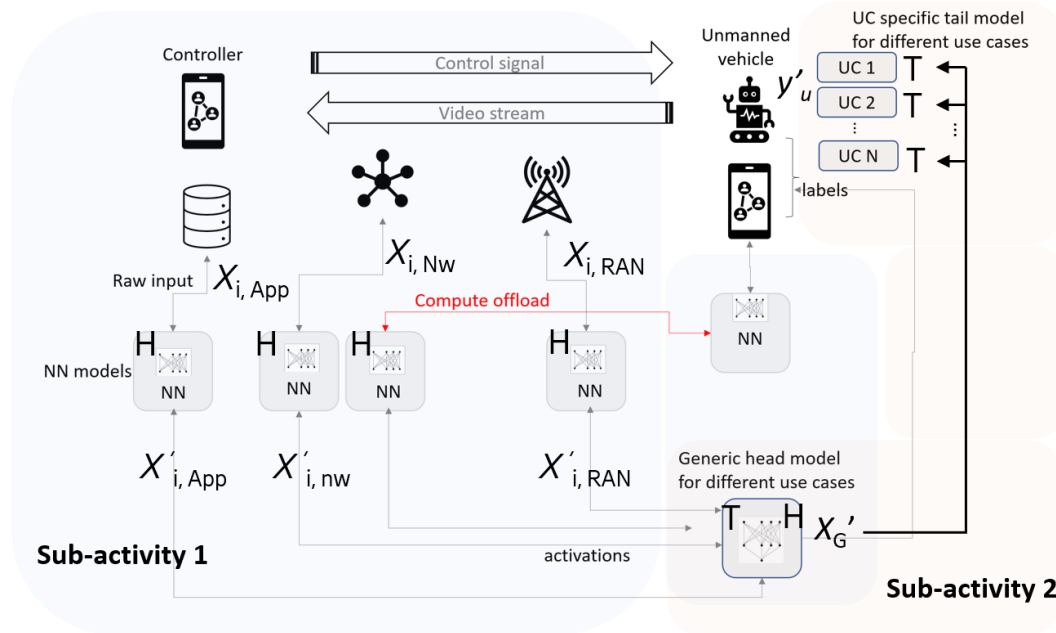


WP3 in Hexa-X-II PoCs

- Hexa-X-II will develop three System-PoCs (A, B, and C) to showcase its ambition and address project objectives
- The three System-PoCs will follow an incremental approach, encompassing more Component PoCs as the developments evolve
- WP3 contributes to System-PoC B via:
 - Network architecture and transformation enablers (all WP3 tasks)
 - Trustworthy flexible topologies-related enablers (T3.3)
 - Beyond communication network enablers (T3.4)
- Solution: *“Networks beyond communications” components offer information on computing resources that can be used; Enablers on data-driven network designate communication resources; A trusted flexible topology, involving nodes from various stakeholders and designated resources is established through the corresponding enablers. As an example, a drone is used, offering also computing resources, linked to some “remote” BS.*



Component-PoC#B.2 : Efficient cross-domain distributed model generalization



Sub-activity 3) Multi-domain: Unsupervised domain adaptation

Many methods for domain adaptation are in fact split architectures
Can be used where data collection is not possible in target environment

- Due to limited data collection capabilities
- No training data

Sub-activity 1) Multi-modal(cross-layer) training and inference: Training a split neural network to estimate QoS using multiple cross-layer observations.

Suppose that there are clients $i: \{1..N\}$, and X_i is a data source (e.g., MAC, RRC, system load, user), and there are N different data sources of different attributes. Y is the observed QoS (e.g., throughput or latency).

Goal: train a joint model with H and T , where H is head and T is a tail ML node. where each H trains a $H_i(X_i): X'_i$ for all input i , then $T(X'_1, X'_2, \dots, X'_i): X'_G$.

Sub-activity 2) Multi task training with SL for the cases when the dataset attributes are the same but the final attributes are different.

Suppose u represent user case identifiers, X'_G : common input (e.g., all input representations), y : target use case variables (e.g., throughput, latency). X'_G is common to all use cases, y are different in all use cases.

Goal: Split the model to generic and personalized partitions, such that $H(X'_G): y'_u$, minimize jointly $y_u - y'_u$ for all u .

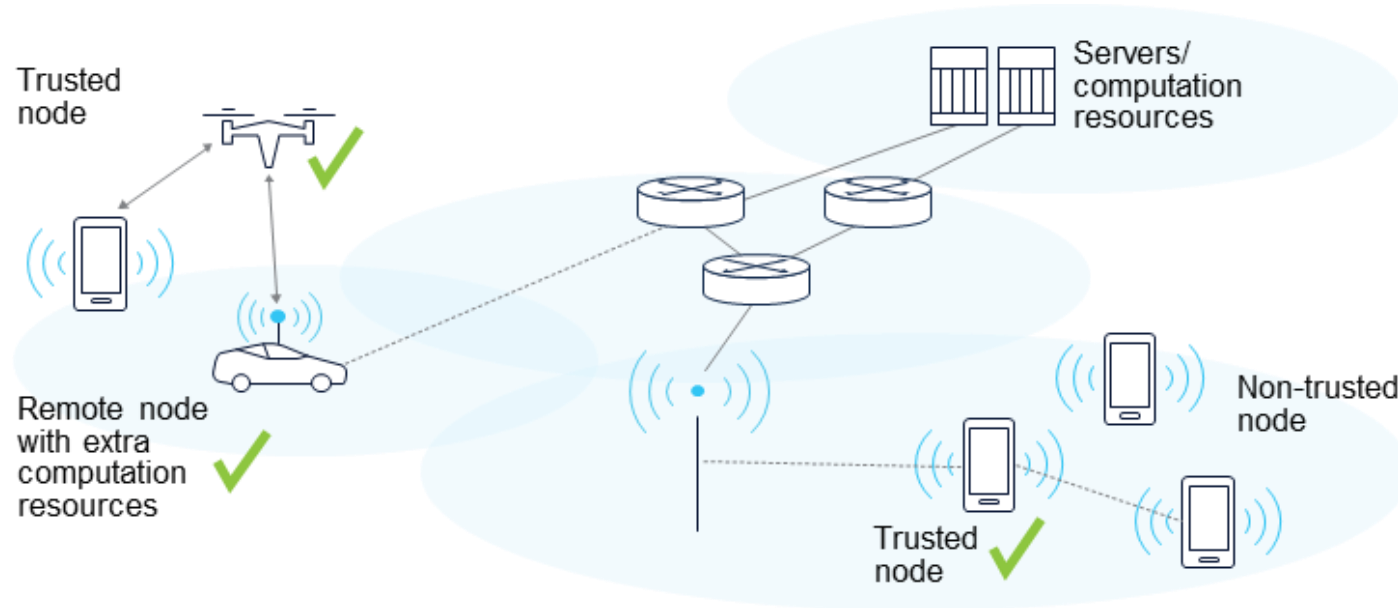
Research Questions:

Multi-modal (cross-layer) learning: How do we train a ML model consisting of inputs from different entities and with different granularity?

Multi-task learning: How do we train minimal number of ML models (with minimal memory/storage requirements) for different use cases and tasks?

Multi-domain adaptation: How do we train a ML model for multiple domains even when labels are not available in some domains?

Component-PoC#B.3 : Trustworthy flexible topologies in 6G, leveraging on “beyond communication” aspects



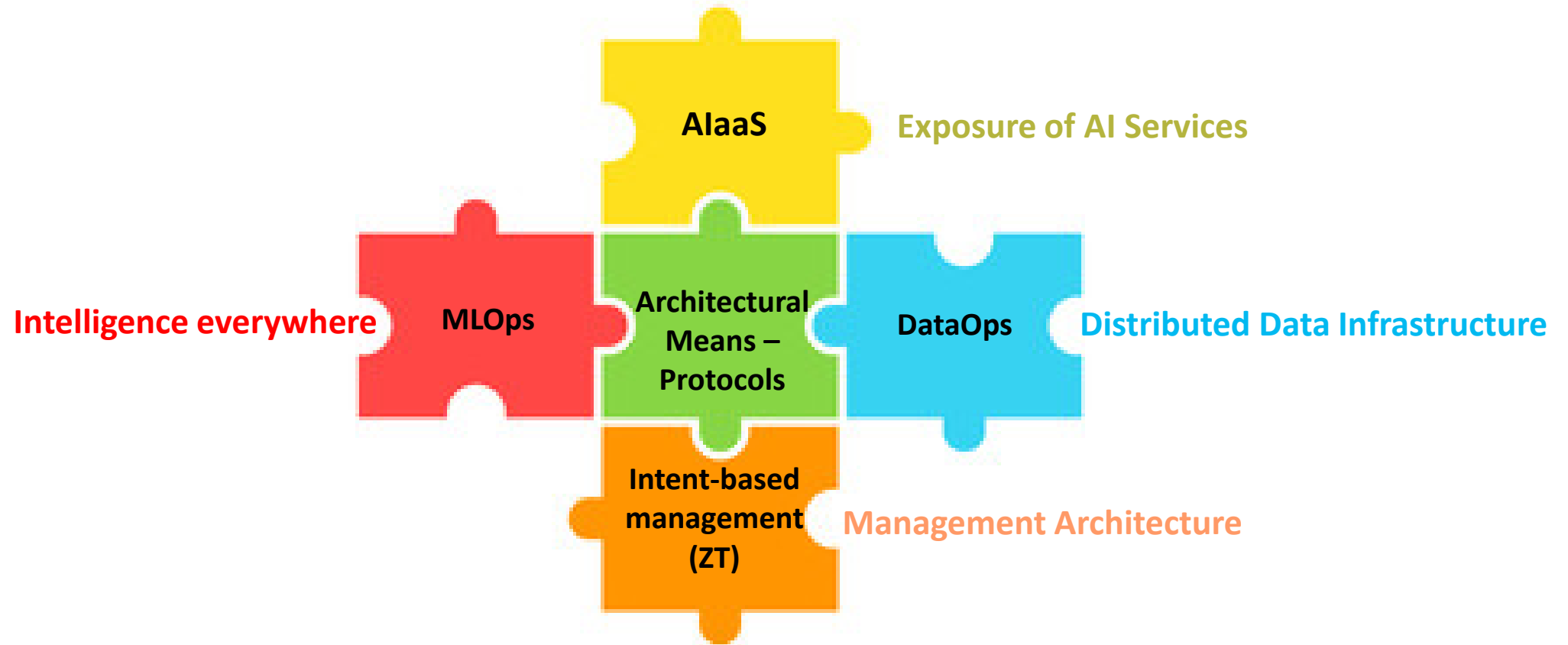
Component-PoC#B.3 targets to integrate aspects related to trustworthy flexible topologies and beyond communication aspects

- Trustworthy flexible topologies (Task 3.3): establish a UAV-based flexible topology offering communication and computing resources, available in near-by edge servers; which nodes must be selected and how to measure the trust levels of the available nodes; which of the collected data should be offloaded towards the edge cloud resources?
- Network beyond communications (Task 3.4): how to efficiently expose and leverage the computing resources available for the specific data-driven services and respective AI/ML workloads that need to be addressed? Which procedures are required to efficiently expose the computing capabilities available in the area, and respectively how to manage the data to be consumed?



AI enablers for data-driven architecture

AI Enablers supporting the Data-Driven Architecture



E3.1.1 AlaaS

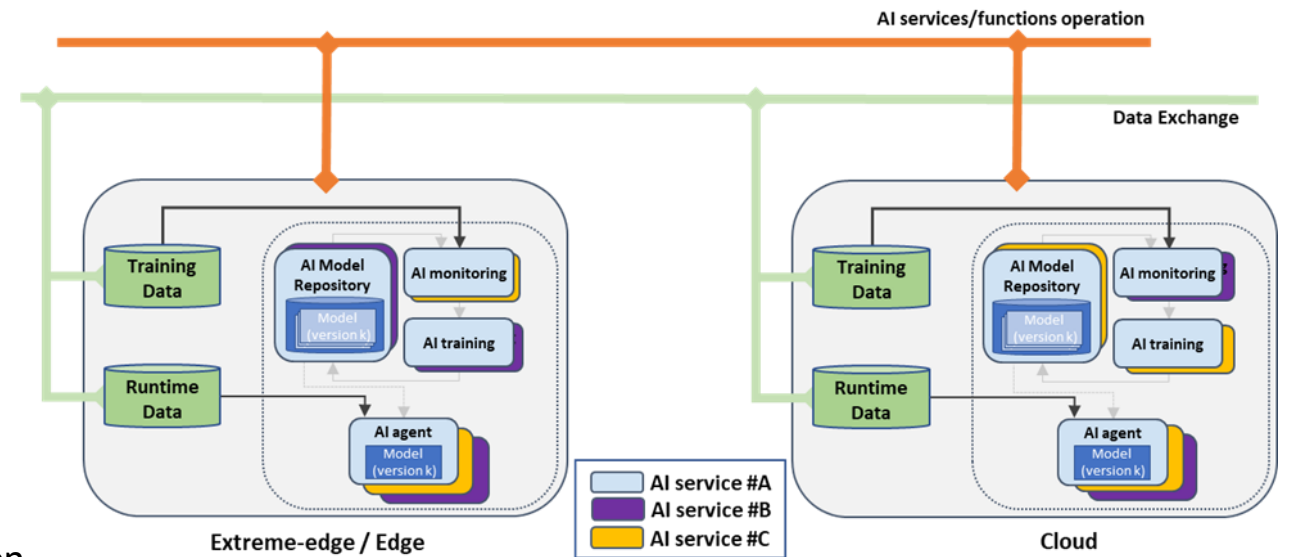


- **Input from studies**

1. AlaaS operation
2. Distributed AI Services
3. Strategies and mechanisms for distributed AI & AlaaS functions management
4. AI Native Architecture

- **Background & Motivation**

- The AlaaS is a comprehensive framework that offers a wide range of AI services as well as personalized inference capabilities to the AI service itself.
 - Apart from pure analytics, prediction, and classification capabilities, the AlaaS framework offers other AI capabilities and services to assist closed-loop network and service automation.
- 6G data driven architecture shall be designed to support different purposes such as exposure of the data driven architecture to verticals.



- **Study Areas**

- AlaaS exposure APIs for AI services and functions operations and management
- Data and ML model management procedures in distributed and cooperative scenarios
 - Data and ML model transfer requirements
 - When and where to train and infer to optimize the amount and type of data exchanged
- Study the trade-offs when considering the extreme-edge/edge/cloud continuum
 - e.g., in terms of data exchange
- AI/ML model orchestration.

E3.1.2 MLOps

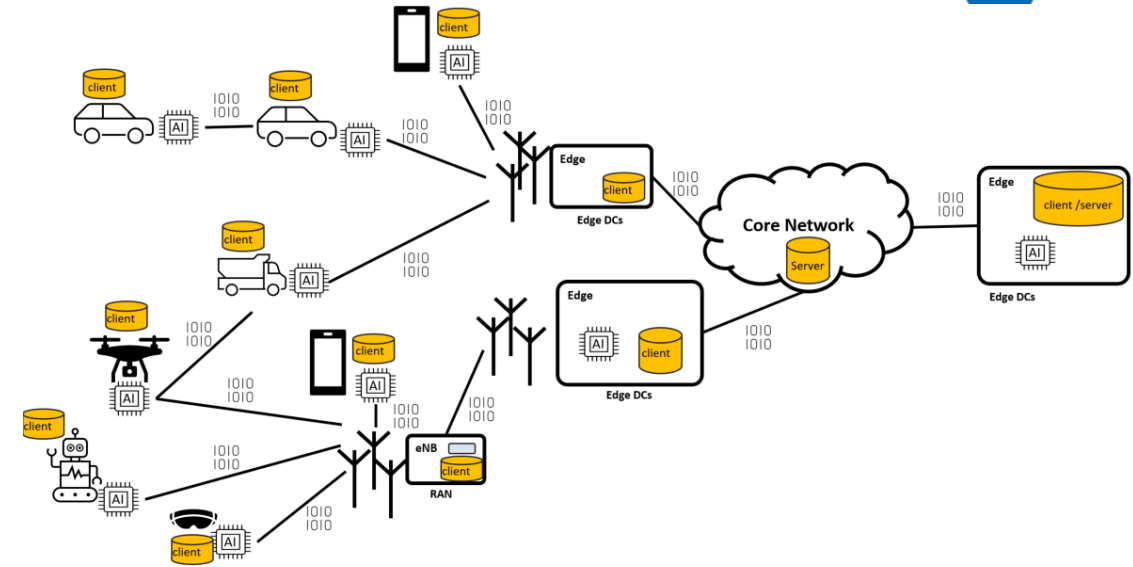


- **Input from studies:**

1. Privacy-aware data collection and learning
2. Sustainable Distributed Model training and Inference
3. AI-native Architecture
4. Distributed AI Services
5. Architectural support for Cooperative Learning
6. Wireless hierarchical federated learning: On the accuracy-energy trade-off
7. Incentive mechanism design for wireless federated learning networks
8. City as an integrated system: federated learning approach between different city verticals
9. Designing a Global AI Model for an E2E 6G NS Instance Employing Distrusted Intelligence Solutions
10. AI-driven coordination of multiple optimization functions
11. Strategies and mechanisms for distributed AI & AlaaS functions management

- **Background and Motivation**

- AI execution environments will be everywhere (e.g., UE, RAN and Core). The increasing number of models will require tools for managing the lifecycle of these AI models considering privacy and trustworthiness aspects.



Decentralized future of network elements and end-devices [ERI-MLOps]

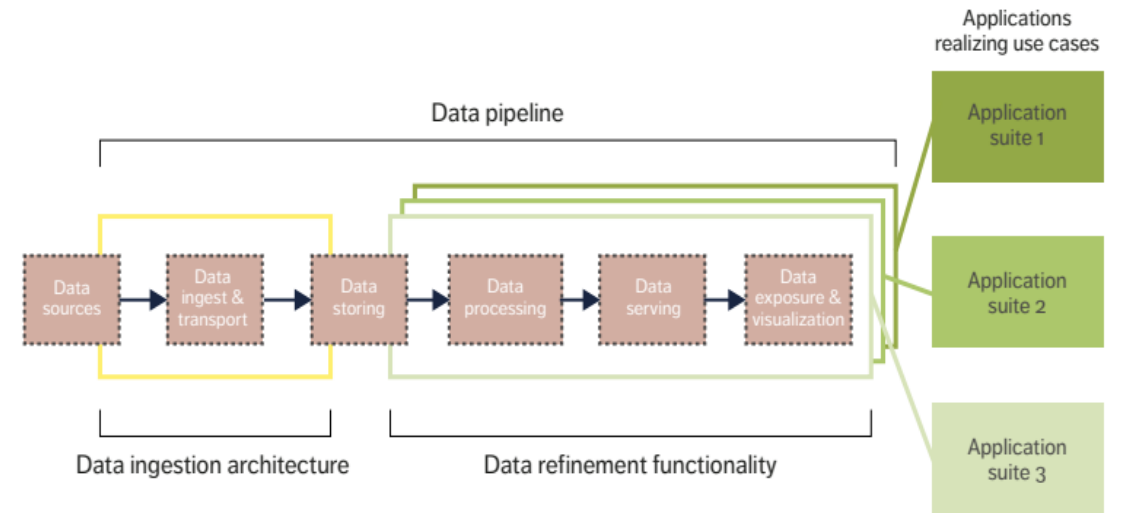
- **Study Areas**

- Collaborative training of a generative or a translator model for only-Network based Inference
- Distributed model and feature selection
- Unsupervised model personalization
- Model evaluation and orchestration mechanisms in efficient distributed model lifecycle management
- Privacy-preserving collection and learning methods: Private Federated Learning and Differential privacy and Coordination
- AI/ML model monitoring
- Distributed AI to coordinate functions

E3.1.3 DataOps



- **Input from studies:**
 1. Sustainable Distributed Model training and Inference
 2. AI-native Architecture
 3. Privacy-aware data collection and learning
 4. Strategies and mechanisms for distributed AI & AlaaS functions management
 5. Distributed AI Services
 6. AI-native Architecture
- **Background and Motivation**
 - The correct functioning of AI is strictly related to data quality. Data shall be delivered, pre-processed, and stored where and when required. This imposes requirements on a flexible data ingestion architecture.
- **Study Areas**
 - Distributed feature selection
 - Privacy-aware data classification
 - Definition and development of data and model management schemes at the edge and extreme-edge to facilitate federated learning
 - Data Collection and pre-processing to be used in AI Orchestrator.

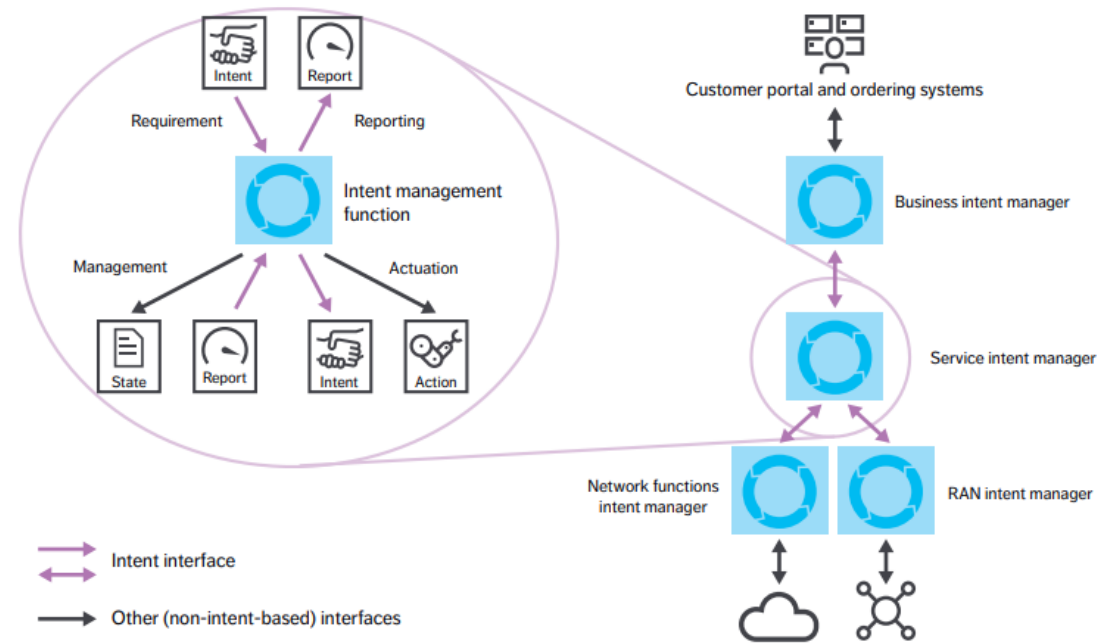


Data pipeline using the harmonized data ingestion architecture [ERI-Data]

E3.1.4 Intent-based Management (Zero-Touch)



- **Input from studies:**
 1. Wireless hierarchical federated learning: On the accuracy-energy trade-off
 2. City as an integrated system: federated learning approach between different city verticals
 3. AI-driven coordination of multiple optimization functions
- **Background and Motivation**
 - Managing intelligence and data will require techniques that simplify manual operations. Zero-touch networks, powered by intent-based management, constitute a paradigm to reduce the complexity arising from the increased adoption of AI models and the number of 6G use cases.
- **Study Areas:**
 - Distributed solutions that allow user self-adaptation
 - AI-based Solutions: Automatic monitoring and management of the city as a whole
 - KPI Monitoring



The intent management function and intent-based networking system stack [ERI-ZTN]

E3.1.5 Architectural Means - Protocols



- **Input from studies:**

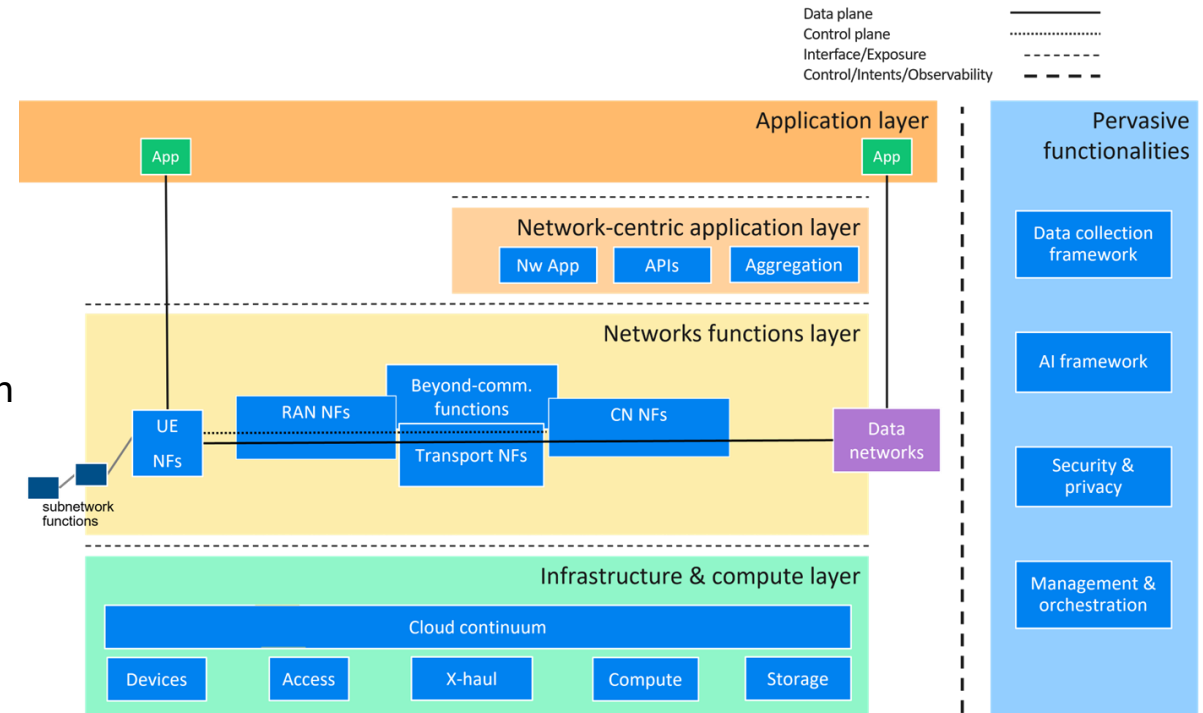
1. Sustainable Distributed Model training and Inference
2. Architectural support for Cooperative Learning
3. AI-native Architecture
4. AI-driven coordination of multiple optimization functions

- **Background and Motivation**

- 6G data-driven architecture will require architectural support that enables communication for AI. This imposes requirements on architectural means and protocols encompassing signalling and ML-model LCM control mechanisms.

- **Study Areas:**

- Split Neural Networks communication protocol
- Discovery procedures and signalling for capability exchange among cooperative cellular nodes.
- Control architecture of cooperative cellular nodes
- MLOps control loop
- Model and data dependency relationship repository
- Modules for Control Loop
- AI-driven coordinator/recommender



Initial 6G E2E system blueprint [HEX23-D21]



Network modularisation

Network modularization enablers and how they fit together.

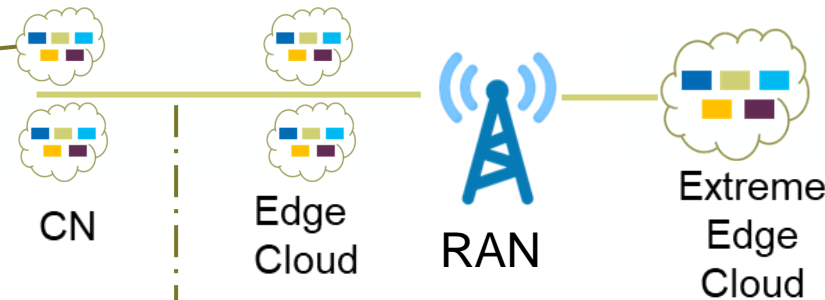
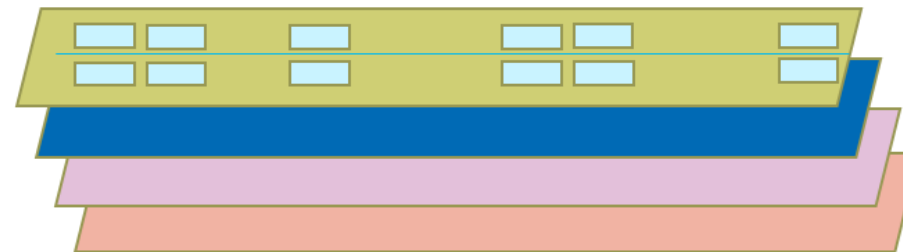
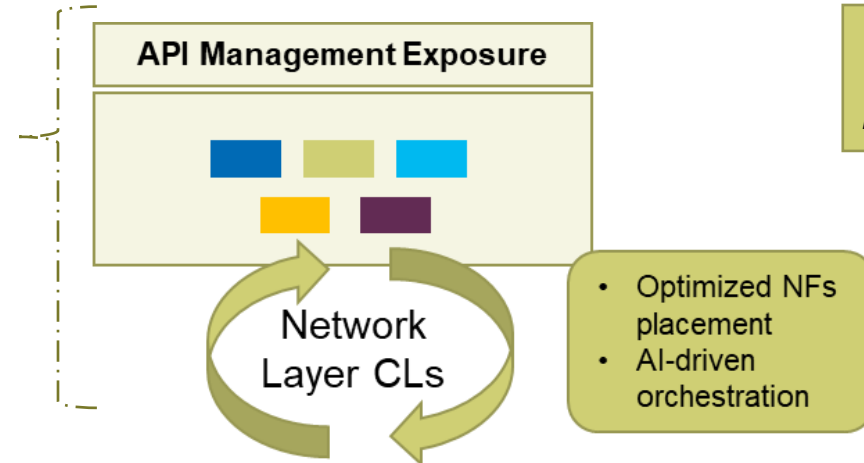
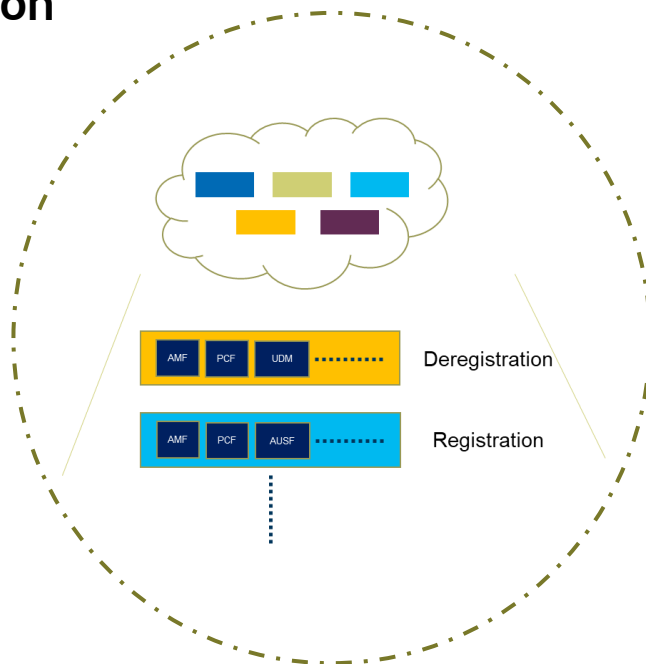
The numbering corresponds to the specific enabler section.



E3.2.4 Network autonomy / Multi-X orchestration

E3.2.5 Network Migration

E3.2.1 Optimized network function composition



E3.2.3 Flexible feature development and run-time scalability

E3.2.2 Streamlined network function interfaces & interaction



E3.2.1 Optimized network function composition

• Background and Motivation

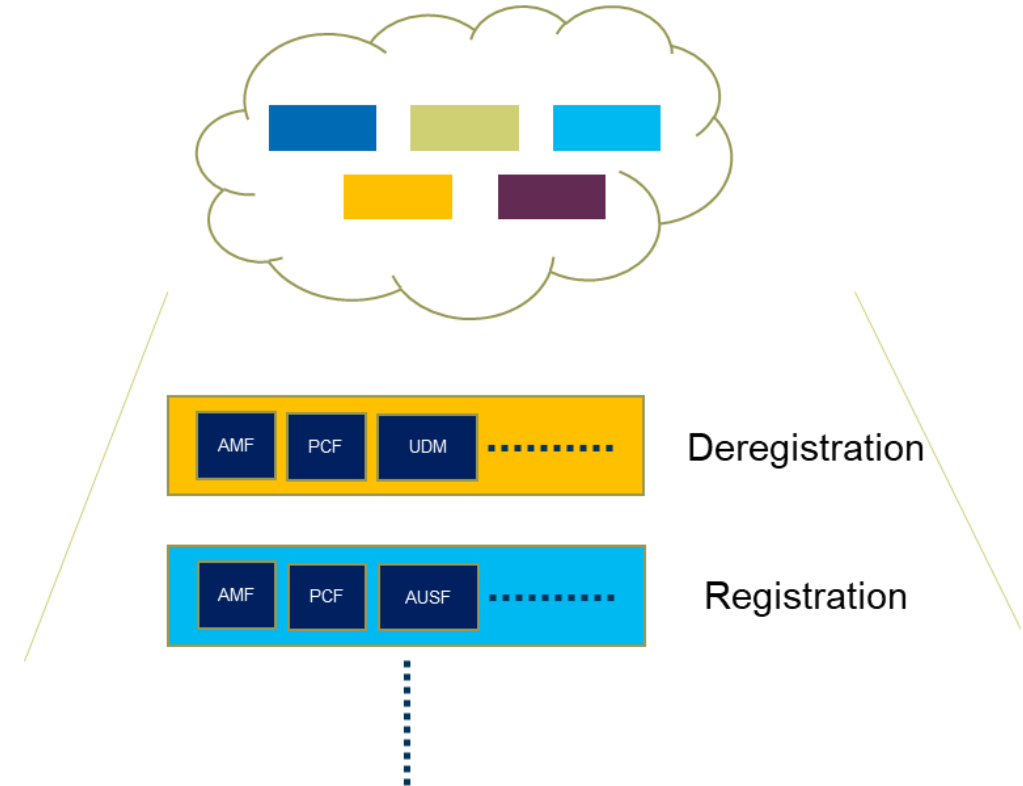
- The 5GC is built upon various network functions (e.g., AMF, UPF) and their interactions to complete pre-defined procedures. The inter function dependencies can cause high signaling costs and latency.
- Modular design minimizes the dependencies between different modules while the relevance of the NF functionalities within a module is maximized. It is possible to further optimize the modules for specific services or deployment options to meet specific KPI targets.
- Contributions to 6G networks include increased flexibility, optimized signaling, and efficient resource usage.

• Study Areas

- Analysing the 5G network function composition
- Exploring the trade-off between network function granularity and performance.
- Mapping the KPIs of the different deployment options and application requirements.
- Investigating the options to optimize the network functionality via network function (de)composition

• Input from studies

1. Procedure-based functional (de)composition for 5G core networks
2. CN-RAN Refactoring
3. Optimised composition and placement of 6GC functions
4. Efficient signaling - separation of concerns
5. Network modularization over the cloud continuum



A possible way to optimize the network function composition

E3.2.2 Streamlined network function interfaces & interaction



• Background and Motivation

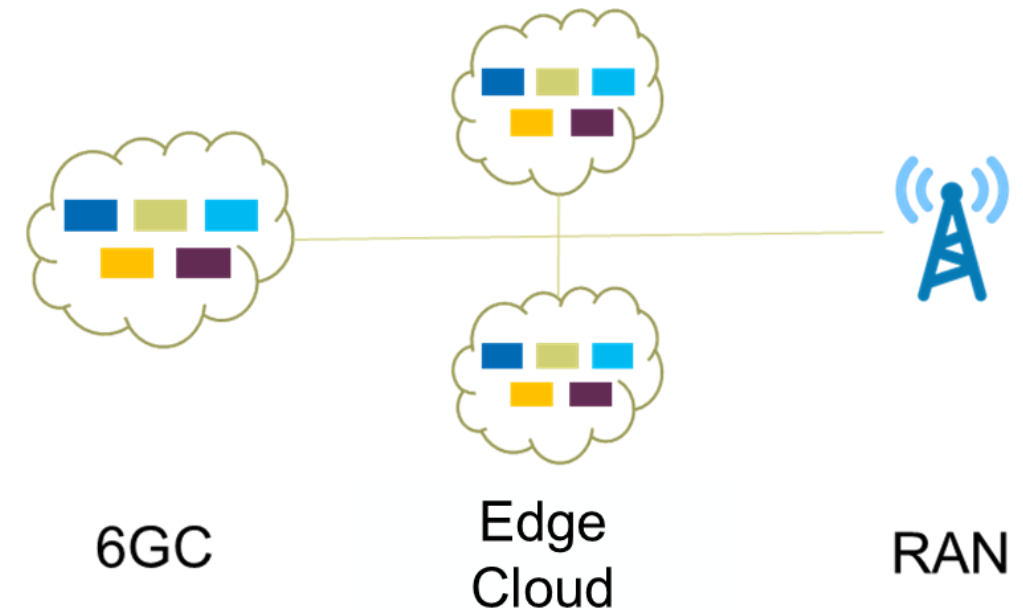
- 6G aims at extending various use cases. The network modules and their interfaces need to support the coexistence of these use cases as well as the related services.
- A new decomposition of the core functionalities may be needed to optimize the signaling while providing the necessary flexibility.
- A dynamic 6GC decomposition and the necessary interface design to support both distributed and centralized deployments.

• Study Areas

- Architectural modules and their external interfaces for versatile use cases, e.g.,
- Extreme edge
- Cell-free experience, etc.
- Distributed to centralized cloud deployments considering resiliency and KPIs

• Input from studies

1. CN-RAN Refactoring
2. Network Modularisation in Hybrid 6G-quantum Architecture
3. Data centric SBA for Edge Native 6G Networking
4. Efficient signaling - separation of concerns
5. Optimised composition and placement of 6GC functions
6. Flexible UPF design
7. Network functions capability exposure & communication
8. Cell-free mMIMO in disaggregated RAN



Streamlined network function interfaces and interaction through cloud, edge and RAN continuum

E3.2.3 Flexible feature development and run-time scalability



- **Background and Motivation**

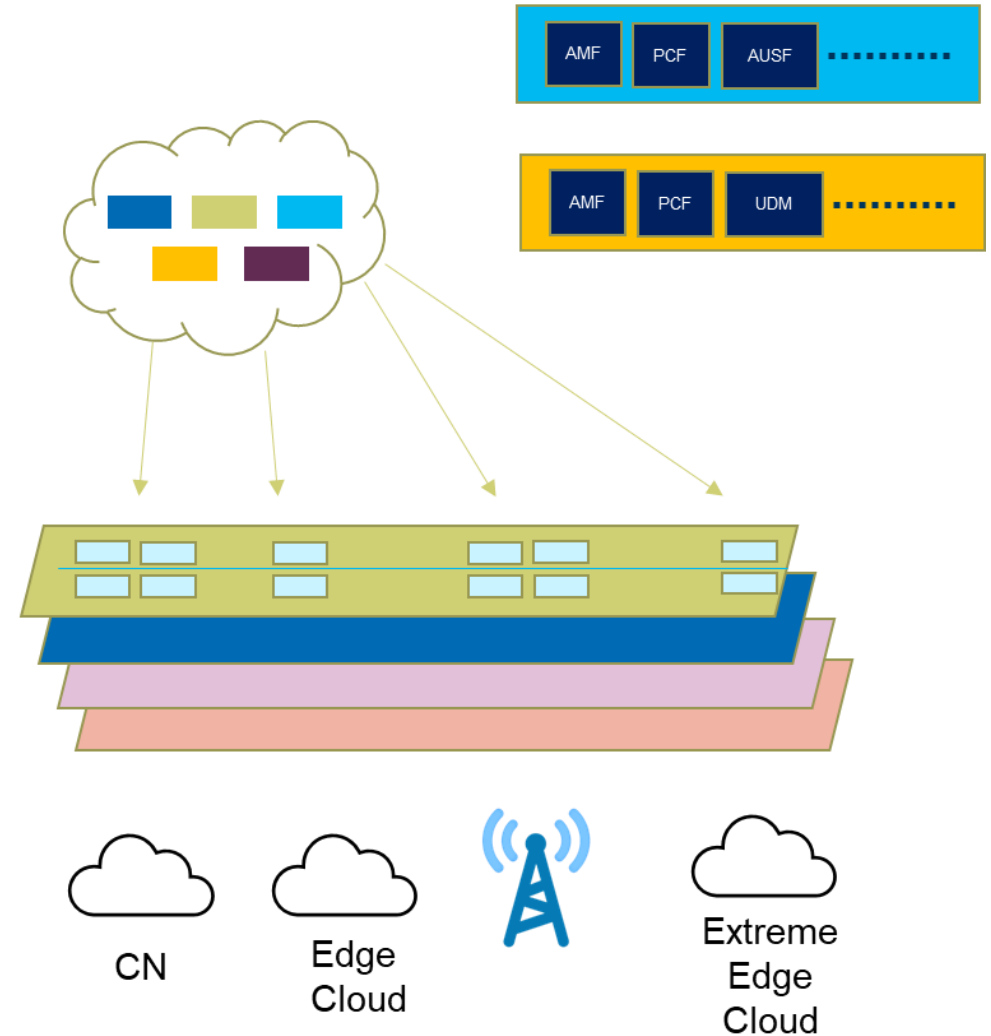
- Exploring the possible enhancements to the E2E modularization (e.g., network slicing in 5G) to optimize the functionality composition to match 6G use cases and performance.

- **Study Areas**

- Analysis of modularity with an E2E disaggregation perspective
- Enhanced network slicing, i.e., the minimum set of functionality per slice/per tenant

- **Input from studies**

1. CN-RAN Refactoring
2. Flexible UPF design
3. Slice as meta module to aggregate separate modules
4. Split management of network slices
5. Cell-free mMIMO in disaggregated RAN



Enhancing the network slicing with network modularity to support flexible feature development and run-time scalability with modular network functionality

E3.2.4 Network autonomy & Multi-X orchestration



• Background and Motivation

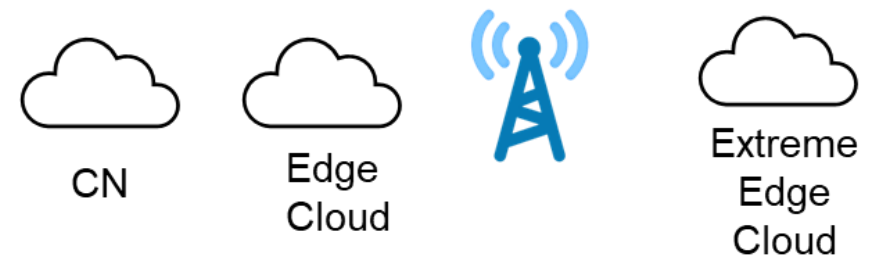
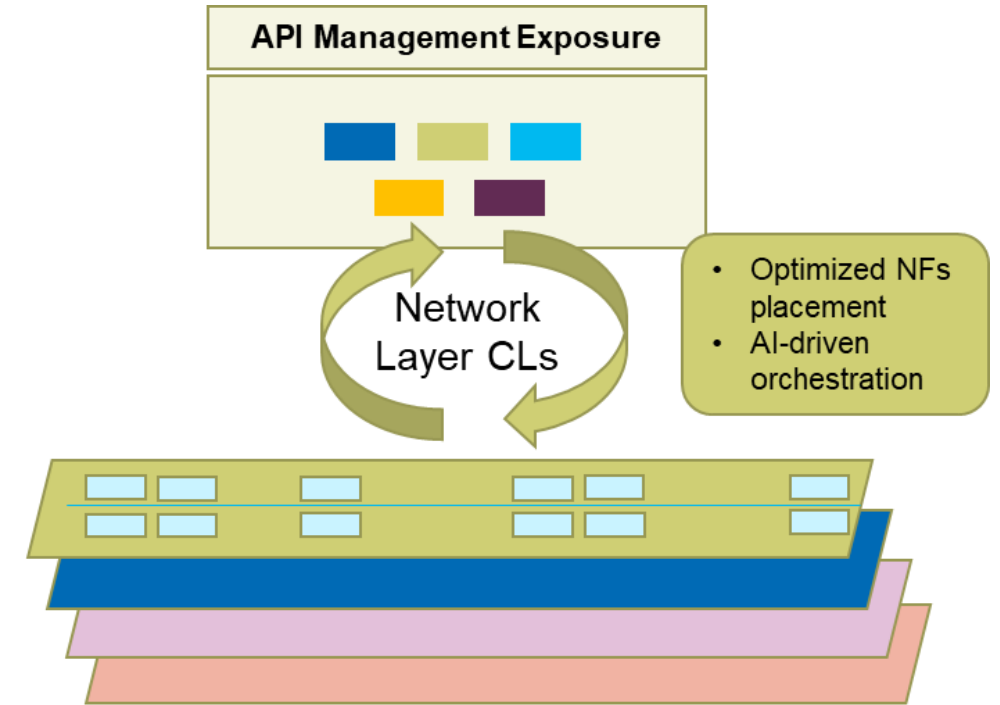
- In 5G, network slicing was a key enabler to facilitate the co-existence of various use cases with demanding and often conflicting requirements. The management and orchestration is built upon open loop slice configurations which often result in low resource utilization.
- The slice-based data or management operations are not structured and strongly biased towards Operations/Business Support Solutions (OSS/BSS)
- In 6G, the communication interfaces/APIs between slices and the respective inter-slice/inter-NF capability exposures need to be clearly defined

• Study Areas

- Exploring the extend of network functions to control a network slice.
- Analysing how these functions may vary depending on the level of control

• Input from studies

1. Slice as meta module to aggregate separate modules
2. CN-RAN Refactoring
3. Split management of network slices
4. Network modularization over the cloud continuum
5. Network functions capability exposure & communication



Network autonomy and multi-x orchestration (X indicating domain, plane or stakeholder etc.)

E3.2.5 Network Migration



• Background and Motivation

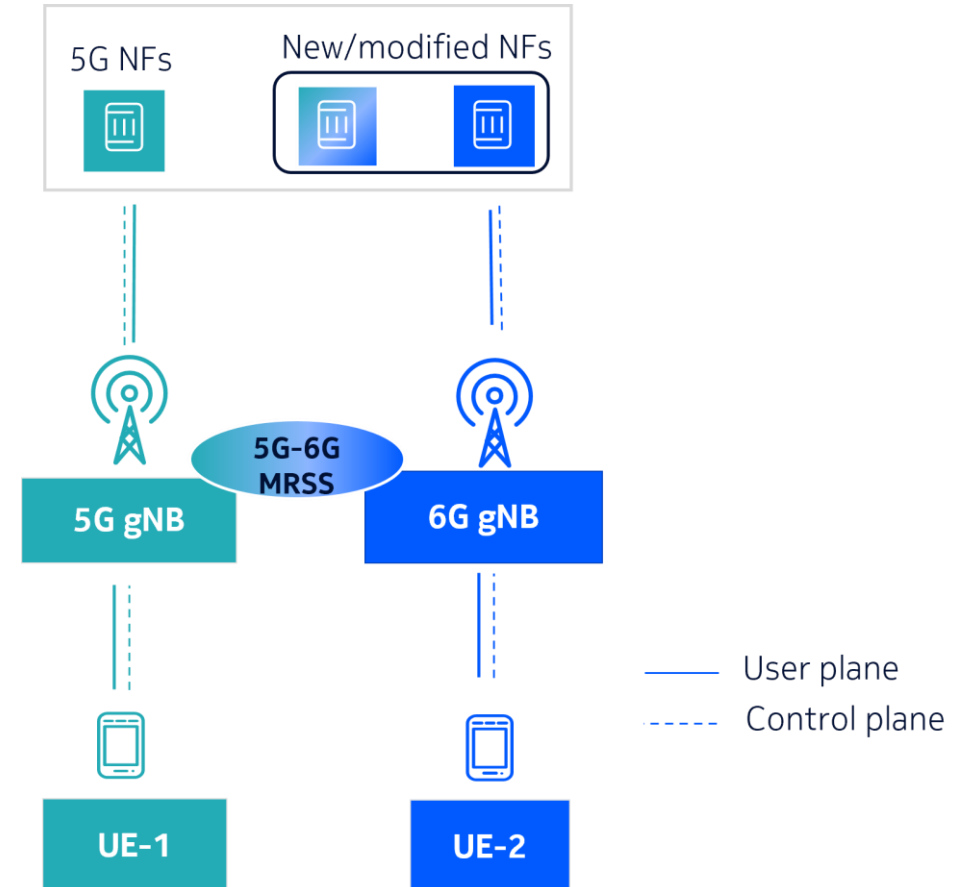
- Multiple non-standalone (NSA) options were considered for the migration from 4G to 5G. This approach increased complexity in network operations since NSA now co-exists with standalone architecture (SA)/native 5G architecture.
- Migration from 5G to 6G should not repeat the complications that are faced in the previous generations. We should simplify the migration. To perform this transition as efficiently as possible 6G should take this into consideration from the beginning: how to migrate from 5G to 6G.

• Study Areas

- A study on the migration options from 5G to 6G
 - Determining what role NSA architectures can play
 - Analysing Multi-RAT Spectrum Sharing (MRSS) with a standalone (SA) deployment
- Evolution of 6G functionality
 - Gradual introduction/modification of new functionalities
 - Investigating which 6G functionality would justify non-backwards compatible changes

• Input from studies

1. 5G-6G MRSS and 6G RAN coordination
2. Evolved core network and lower layer split



A possible 5G to 6G migration path, with new CN NFs and use of spectrum sharing (MRSS) between 5G and 6G



Architectural enablers for new access and flexible topologies

E3.3.1 Network of Networks

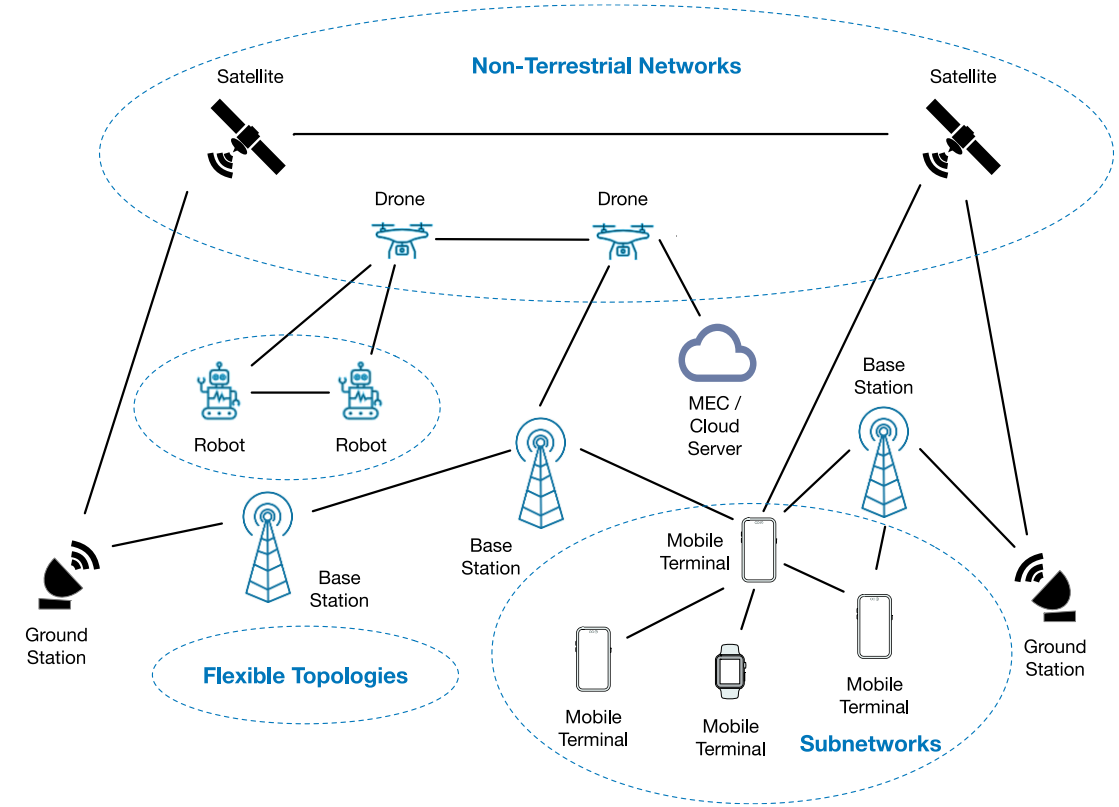


- **Background and Motivation**

- Network of networks enables the integration of multiple subnetworks, including terrestrial and non-terrestrial networks in order to create a seamless and ubiquitous communication system
- This enabler can contribute to the goals of 6G networks, such as extreme coverage, reduced complexity, increased reliability and more efficient management of network resources
- Fulfills 6G design principle #3 (Flexibility to different network scenarios)

- **Study Areas**

- Study the architecture of flexible topologies, such as subnetworks formed by multiple user-owned and/or NW-owned nodes (terrestrial, aerial, non-terrestrial)
- Research the new roles and responsibilities of the nodes as well as the coordination between the nodes
- Investigate architectures with efficient inter-satellite-link hops to enable true global service coverage
- Design unified decision-making and resource allocation frameworks for the subnetworks
- Study management procedures for trustworthy flexible topologies
- Develop data-driven ML tools that predict future coverage developments and include proposed solutions in the prediction models



- **Input from studies:**

1. Subnetworks: Architecture, new roles and responsibilities of the nodes, coordination between nodes
2. NTN architecture and global coverage
3. Digital continuum: Architecture design and decision making
4. Large-scale coverage prediction tools for flexible topologies
5. Trustworthy, Flexible, Unstructured Networks

E3.3.2 Multi-Connectivity



• Background and Motivation

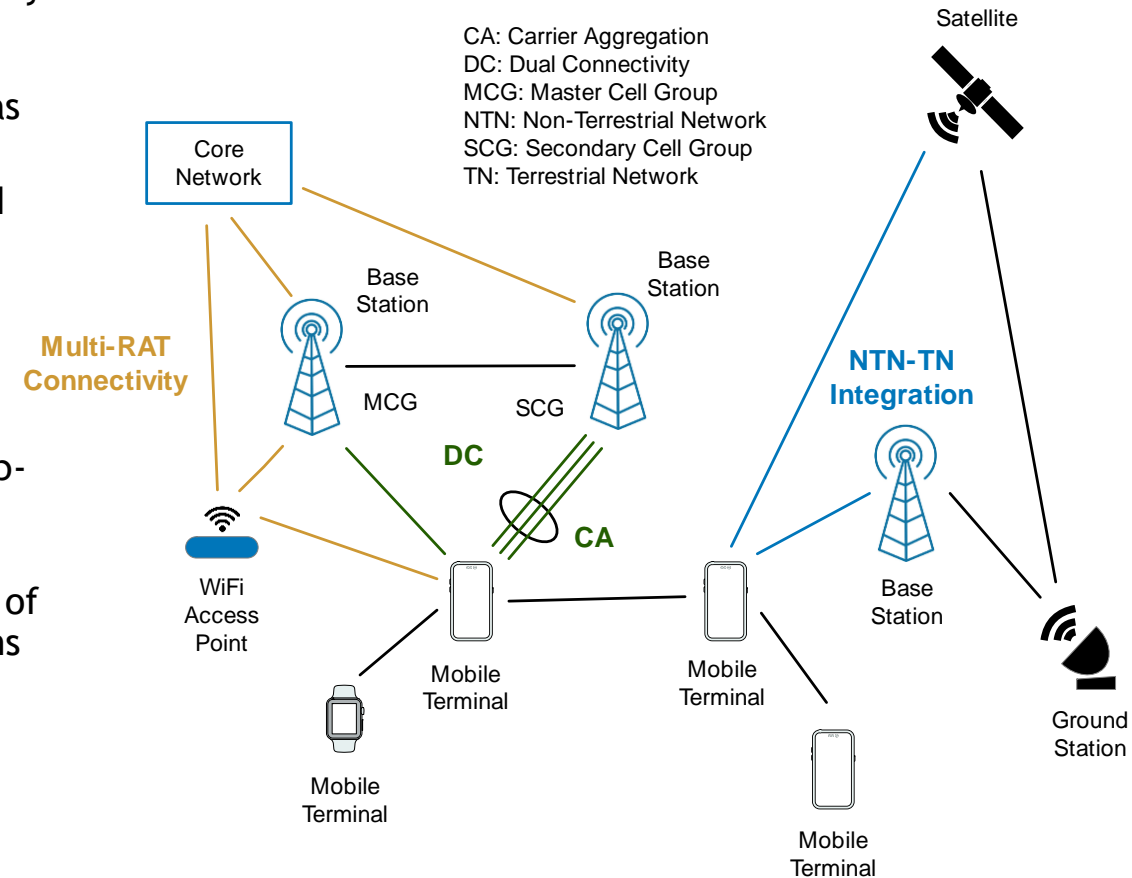
- Multi-connectivity enables the configuration and use of multiple frequency ranges by different physically separated nodes, the aggregation of different radio access technologies, carriers, and access networks, the communication with both terrestrial and non-terrestrial nodes, as well as the integration of subnetworks to the parent network
- Contributions to 6G networks goals include extreme coverage, increased reliability, increased flexibility
- Fulfills 6G design principle #5 (Resilience and availability) and #3 (Flexibility to different network scenarios)

• Study Areas

- Investigate different types of multi-connectivity, such as cellular and sub-networks, cellular and different radio access technologies, dual connectivity to terrestrial and non-terrestrial nodes
- Design a new multi-connectivity solution combining the positive aspects of the current Dual-Connectivity (DC) and Carrier Aggregation (CA) solutions
- Study aggregation of different radio access networks on different levels (e.g., CN, RAN) and in a connectivity domain-abstracted manner

• Input from studies:

1. Evolution of multi-connectivity
2. 6G multi-connectivity: Combining Carrier Aggregation and Dual Connectivity
3. Abstracted approach to multi-connectivity
4. NTN-TN integration and global coverage



E3.3.3 E2E context awareness management

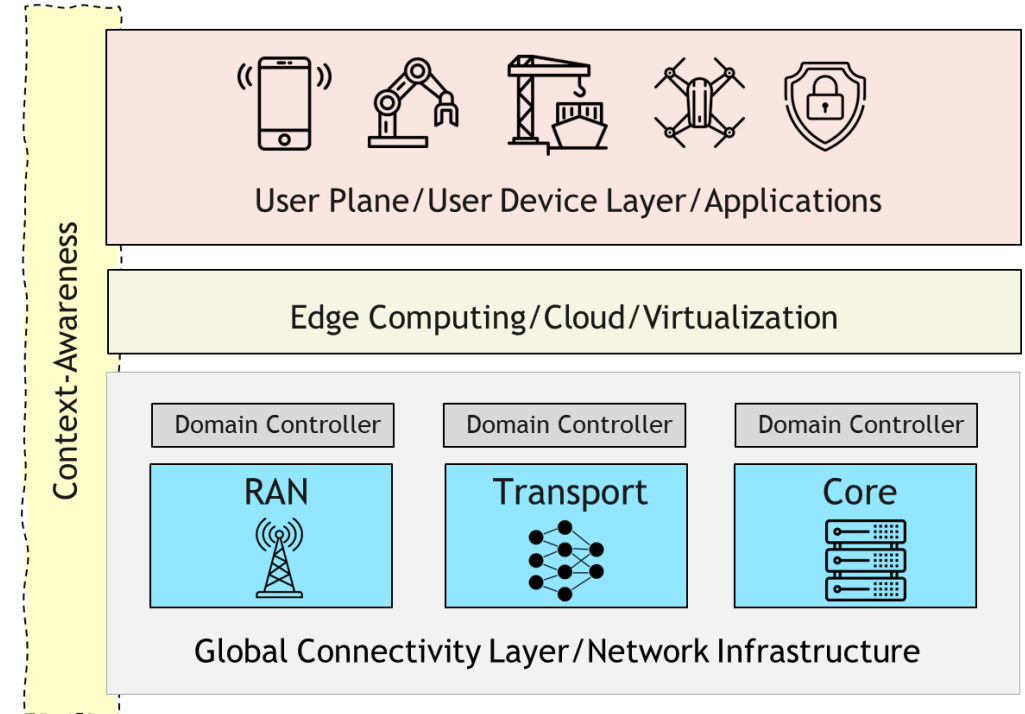


• Background and Motivation

- The concept of **context awareness management** is focused on managing a network so that it can adapt dynamically to the context (e.g., user requirements, layer specific characteristics), to ensure the expected E2E QoS/QoE with the effective use of the network resources.
- This impacts:
 - The network components (e.g., RAN/CN, transport, applications), including automation mechanisms to facilitate the interaction among the infrastructure components, orchestration and slicing, etc.
 - Mechanisms to enable operation in service and network layers based on mutual awareness on the services/application and infrastructure layer
- Fulfils 6G design principle #3 (Flexibility to different network scenarios)

• Study areas

- To enforce context-aware connectivity, a **resource orchestrator** works with an **abstracted** view of the network resources and triggers the domain controllers to handle the respective infrastructure resources, including the **transport** ones. This facilitates the mapping of QoS among the infrastructure components and optimizes the resource usage.
- The **delayed computing** approach promotes higher network capacity by postponing the completion of specific tasks until the network experiences lower traffic congestion. The goal is to design appropriate incentive mechanisms to motivate users to leverage their delay tolerance and price sensitivity in exchange for reduced computing service cost and to properly schedule the tasks at the edge or at the cloud.
- Study context-aware connectivity for **situational awareness in a highly critical environment**, by leveraging the network infrastructure to deploy automation mechanisms. Also, the goal is to use the acquired understanding and knowledge in order to make predictions of future critical events.
- Consider the semantics of **robot** tasks in mission-critical operations, in order to reduce the network overhead and to allocate edge resources flexibly, ultimately improving the system performance by allowing multiple edge allocations and RAN slices.
- Investigate **flow-aware transport**, which considers the mobility of both ends (transport border nodes) of a path, for edge-based applications and for multi-connectivity, in order to remove the limitations of tunnel-based data flow. With flow-aware transport, data can flow based on scalable, multi-domain SDN, supporting the mobility of flows and AI-driven traffic engineering to optimally use available connectivity resources



• Input from studies:

1. Context aware transport
2. Delayed computing paradigm
3. Context-aware connectivity for maritime ports
4. Context-aware and flexible RAN
5. Tunnel-free User Plane supporting mobility of both ends of path



Networks beyond communications

E3.4.1 Exposure and data management

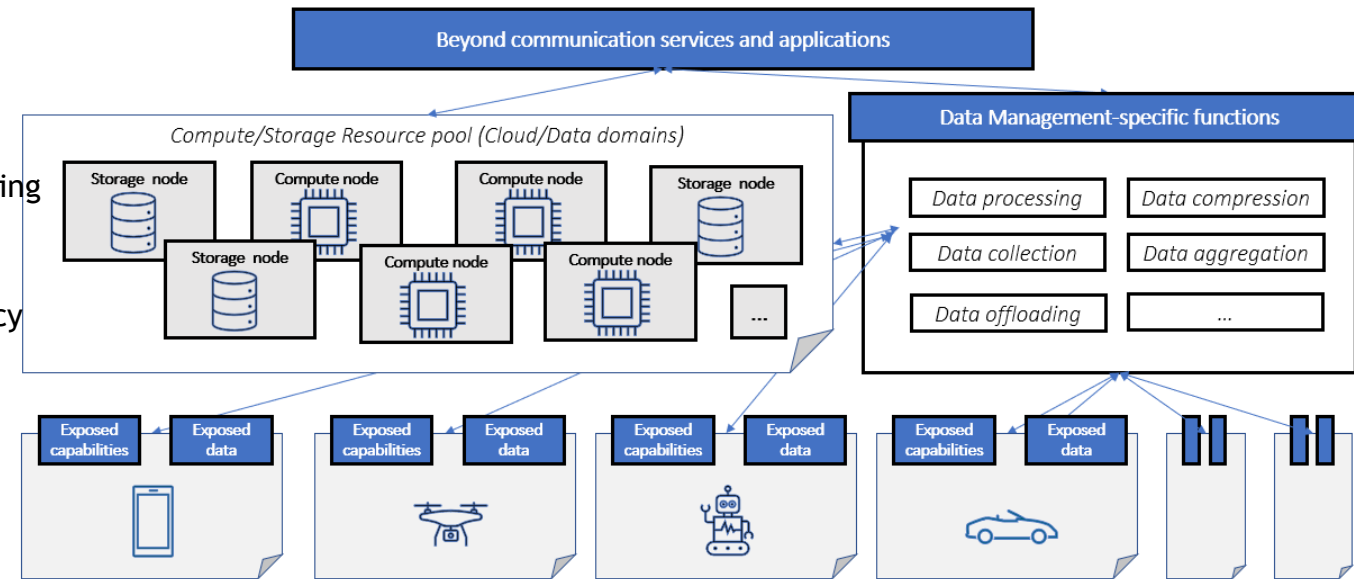


• Input from studies:

1. BCS information exposure and functionality allocation for beyond communication services
2. Defining enhanced functionality for Joint Communication and Sensing services
3. Define interfaces that support data collection, data processing and data distribution, respecting privacy, security, energy-efficiency
4. Discovery, control signaling protocols and procedures for computational offloading
5. New 6G architecture supporting Joint Communication and Sensing use cases
6. Study the requirements on the data volumes (scaling) needed to properly support tracking in chosen use cases

• Background and Motivation

- Emergence of new applications in 6G that require advanced data management capabilities
- Emergence of new devices with diverse capabilities and provided data
- Need for more efficient flexible and scalable data management solutions able to handle massive amounts of data
- Seamless user experience across different applications and services, regardless of the underlying data management system
- New interfaces for data management are needed to enable seamlessly edge computing



• Study Areas

- Device data/capabilities exposure, subject to trust/privacy constraints
- Determining the interfaces for applications that use JCS services
- Design how the data collection, aggregation and labelling to be exposed between the consuming functions in core and in 3rd applications
- Define interfaces that support data collection, data processing and data distribution, respecting privacy, security, energy-efficiency
- Definition of parameters for node discovery / selection
- Scaling the interfaces for supporting JCS

E3.4.2 Protocols, signalling and procedures for BCS

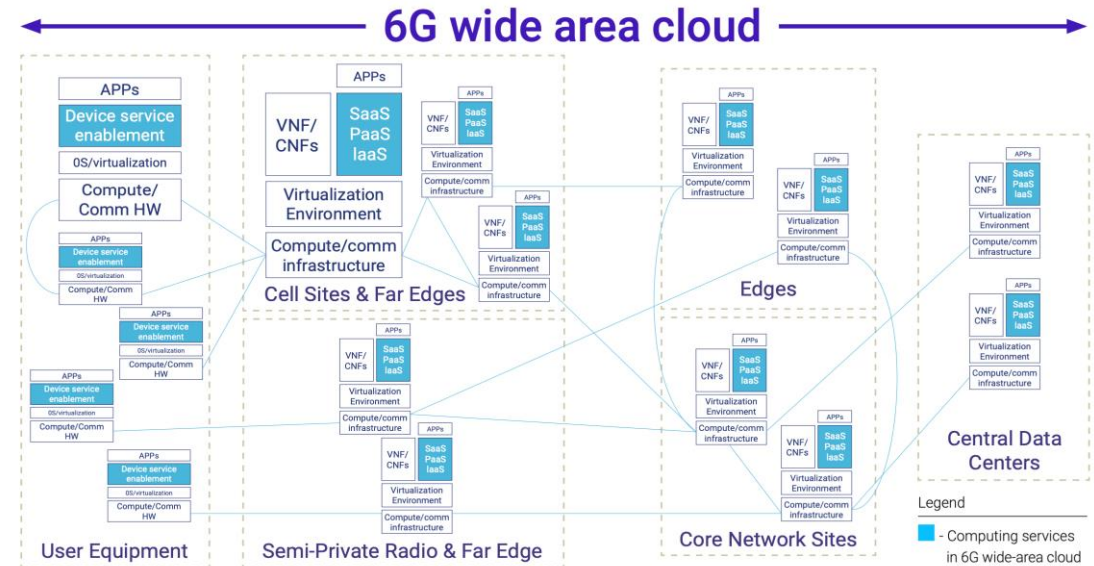


• Input from studies:

1. Distributed compute as a (beyond communication) service
2. Discovery, control signaling protocols and procedures for computational offloading
3. Requirements and protocol scenarios and solutions that enable the introduction of JCAS

• Background and Motivation

- The true convergence of communication and computing is required to support resilient distributed computing
- The discovery of the compute node and corresponding signaling protocols must be carefully designed when introducing the computation in the protocol stack.
- Synchronization and coordination among offloading and computing nodes must be facilitated
- The coordination of communication and computation resources within the network is required to reduce computation times and balance the trade-offs between computation resiliency and communication QoE
- The impact of new sensing services on RAN interfaces and functionality.



• Study Areas

- Optimization of the signaling and procedures for computation
- Characterization of the offloading procedures
- Definition of generic properties of typical offloaded functions
- Degree of offloading
- Deployment (pre-deploy/ad-hoc), scheduling, initiator (device/network)
- Develop requirements and protocol scenarios and solutions that enable the introduction of Joint Communication and Sensing

E3.4.3 Application- and Device-driven optimisation for BCS



- **Input from studies:**

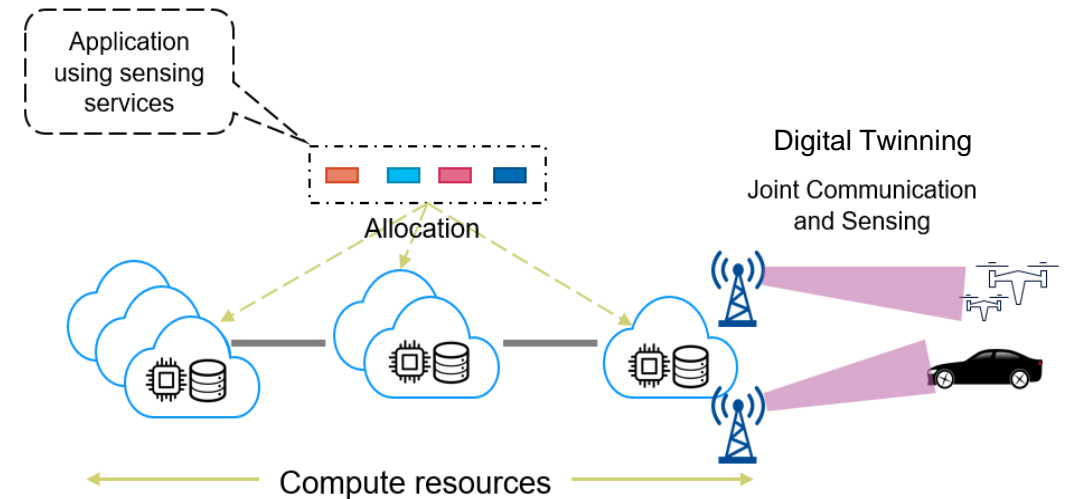
1. BCS information exposure and functionality allocation for beyond communication services
2. Defining enhanced functionality for Joint Communication and Sensing (JCAS) services
3. Quantum-enhanced 6G Communication and Sensing
4. Solve the wide range of devices supported by the network, both from connectivity as well as from a mobility point of view

- **Background and Motivation**

- JCAS is currently associated with radio level functionality. The potential of JCAS is much wider covering applications that use JCAS services.
- Several applications for supporting verticals such as industry 4.0, will also need to satisfy certain QoS/QoE, as well as contradicting requirements for communication, sensing and computation vs sustainability and energy efficiency
- Such applications have their own requirements.
- QoS/QoE for BCS and the contradicting requirements for Communications / Sensing, along with trade-offs like computation vs sustainability and energy efficiency

- **Study Areas**

- Defining the requirements associated to applications using JCAS, Digital Twinning, etc.
- Analysing how applications that use sensing functionality, or applications supporting verticals such as Industry 4.0 (Digital Twinning) should be allocated for BCS.



E3.4.5 Enhancing JCAS capabilities



- **Input from studies:**
 1. Integration of quantum communication-sensing for enhancing 6G beyond communication capabilities
 2. Indoor mapping using mmWave WiFi C&S
- **Background and Motivation**
 - The inclusion of sensing capabilities in a wireless communication networks is a very promising area that can find applicability in many new use cases (e.g., Human Localization and Tracing, SLAM, 3D vision)
 - On one side, many conceptual studies, e.g., [WSL+21], have identified JCAS as the main 6G enabler, but none of them elaborates on functional solutions for JCAS.
 - In addition, the massive distributed communication and sensing can raise significant limitations for classical technologies because of communication overhead and usage of resources.
- **Study areas**
 - Quantum communications for JCAS
 - mmWave-based JCAS for SLAM



Figure from: [JSAC-1]



Virtualisation and Cloud transformation

E3.5.1 Integration and orchestration of computing continuum resources



- **Input from studies**

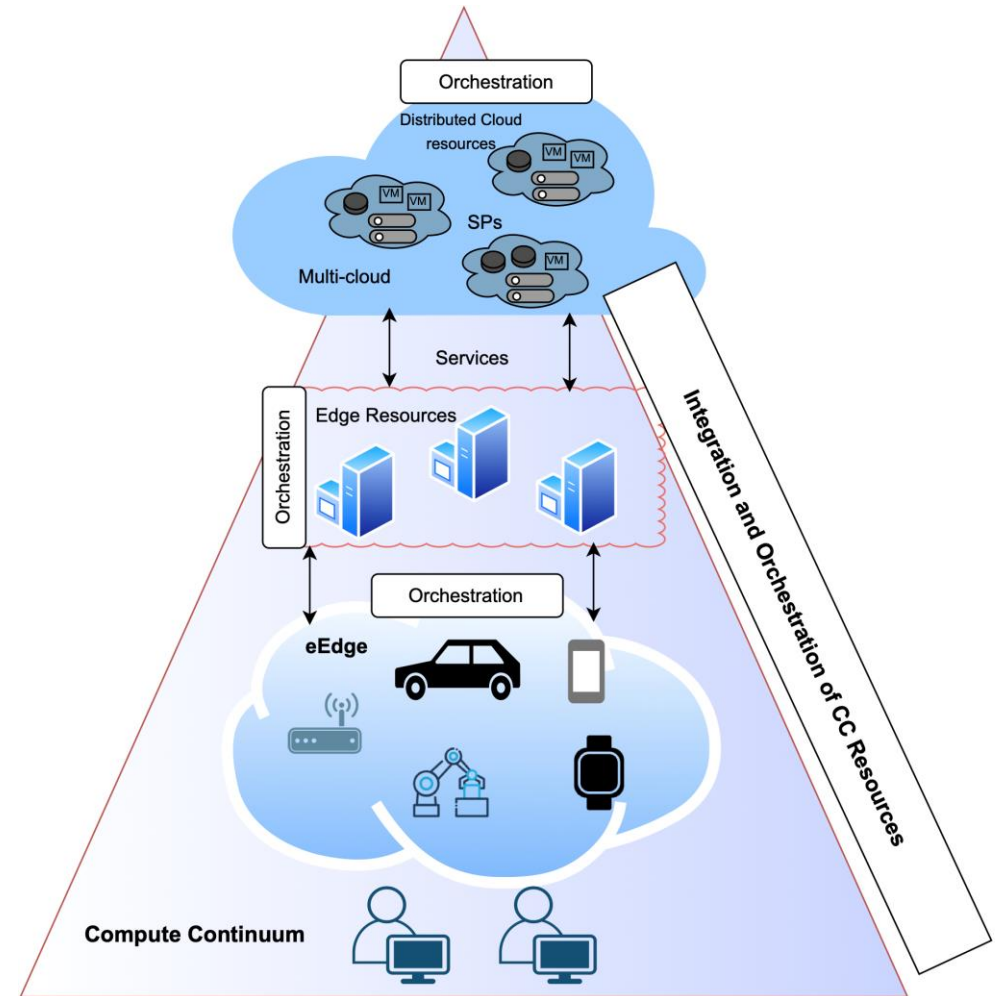
1. Continuum Management & Orchestration
2. Extensions of ETSI MEC framework in constrained devices
3. Decentralized compute continuum smart management

- **Background and Motivation**

- Future 6G networks will consider computing capabilities across the full network, from the extreme edge (including the UE) to Telco grade clouds (CC, Compute Continuum).
- Management of the resources in the CC and the orchestration of services across the integrated computing infrastructure is key.
- The CC poses challenges not considered in current architectures, such as volatility of resources, capacity heterogeneity, heterogeneous APIs, different administrative domains. etc.

- **Study Areas**

- Discovery and monitoring of CC resources
- Integration of resources into a computing continuum
- Serverless computing integration in e2e orchestration approaches
- Addressing volatility of resources
- Integration of constrained devices in edge architectures



The three layer compute continuum: xEdge, Edge and Cloud resources

E3.5.2 Multi-domain/Multi-cloud federation



- **Input from studies**

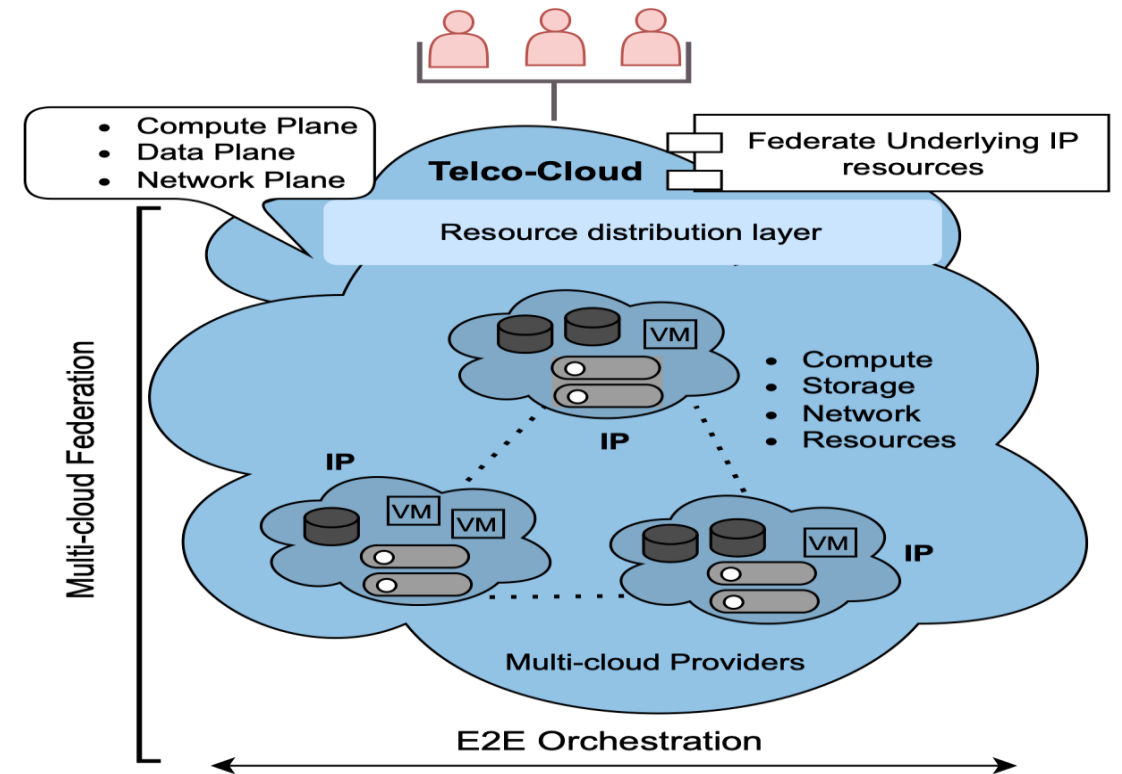
1. Multi-cloud orchestration in federation scenarios
2. Multiple providers in the cloud continuum concept
3. Multi-cloud federation in data centres

- **Background and Motivation**

- Different Telco-Cloud Providers are offering Compute, Storage and Network resources as a service
- 6G networks will require of on-demand federation of such cloud resources located in different clouds to allow them to be more cost-effective and adaptable to user demands
- Examples of KPIs affected by this enabler are coverage extension, service creation time, performance improvements in terms of latency and throughput, etc.
- An end-to-end management and orchestration framework for multi-domain cloud federation is needed.

- **Study areas**

- Unification of existing domain specific orchestration frameworks into an e2e federated architecture.
- Definition of federation interfaces to realize the vision of having a single Telco Cloud based on resources from different infrastructure providers.
- Clear requirements towards the end-to-end orchestration and management plane for 6G networks



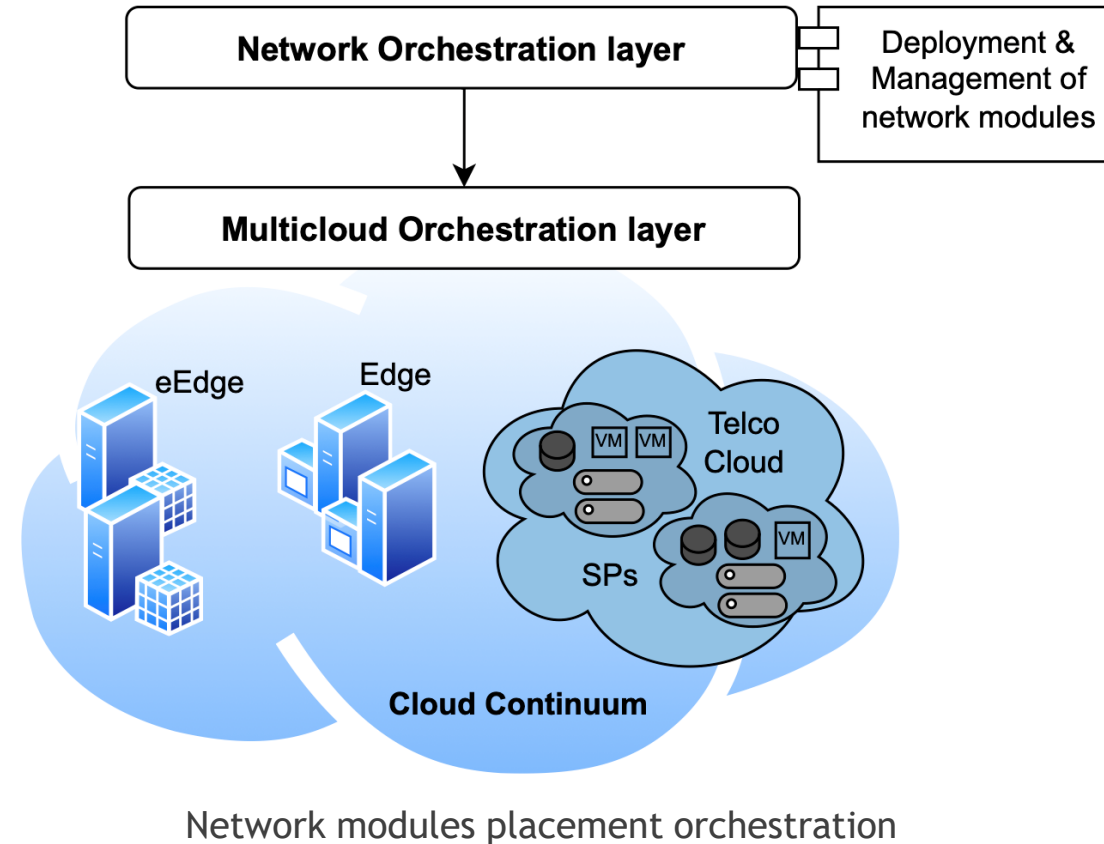
Multi-domain/Multi-cloud Integration model

- Analysis of possible federations to answer distinct applications
- Analysis of business relations including extreme edge resource integration

E3.5.3 Network modules placement



- **Input from studies:**
 1. ETSI MEC deployment in constrained devices
 2. Network module placement across cloud continuum
- **Background and Motivation**
 - 6G network architecture is assumed to be more modular than 5G. This will be achieved by reconsidering the modules involved in the different network operations and creating atomic functions that can be used to compose services.
 - These new functions or modules, will run on top of the CC, spanning resources across a variety of locations and capabilities.
 - The function placement in the CC is challenging considering the heterogeneity and volatility of the Edge and Extrema Edge resources.
 - The correct function placement of the new modules with real-time requirements is critical for the correct operation of the network.
- **Study Areas**
 - Definition of APIs to expose capabilities of heterogeneous computing resources in the CC.
 - Develop techniques to optimize the placement of network functions of modules based on their real-time requirements and performance trade-offs.
 - As an example, study the new modular split of ETSI MEC to consider its deployment in extreme edge devices.



E3.5.4 Cloud Transformation in 6G-quantum architecture

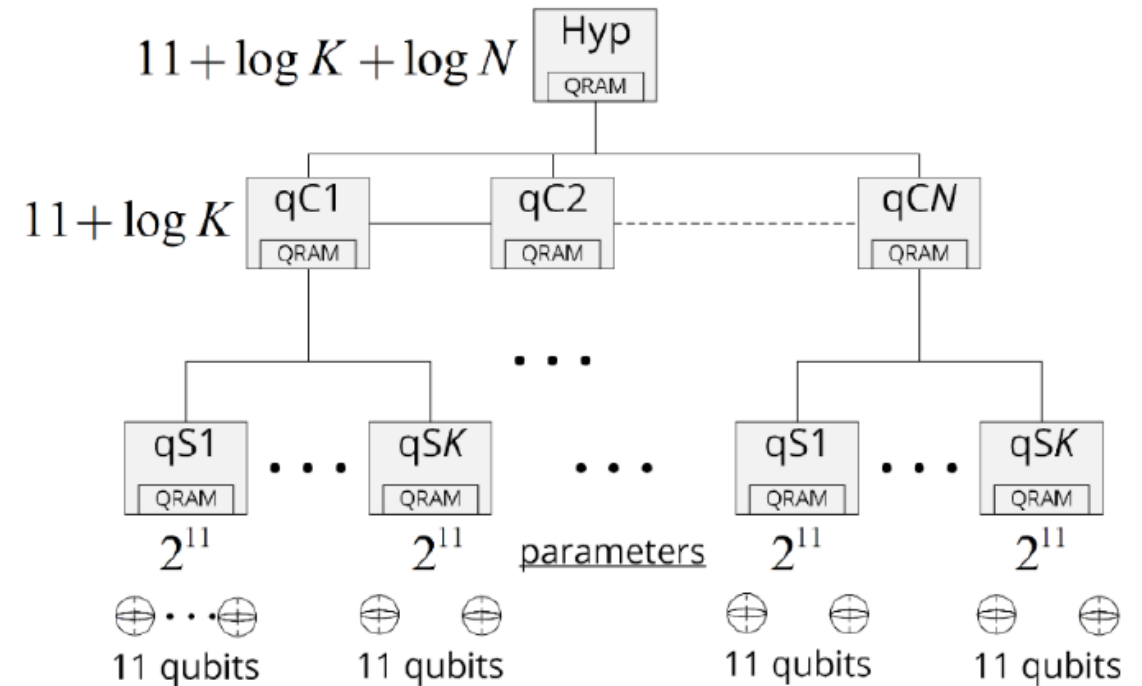


- **Background and Motivation**

- The 6G cloud supporting the softwarized network continuum will imply the explosion of network control traffic.
- Quantum technologies and a 6G-quantum network architecture can improve the optimal use of resources and the intelligent network management by also reducing the problem of big data analytics (e.g., mining, classification, transmission). By using lossy encoding of data in quantum bits and adapting the cloud with algorithms to process them, it is possible to reduce significantly the load of data mining procedures.

- **Study area:**

- Integration of quantum technologies with cloud and softwarized continuum for the cloud transformation in future 6G-quantum network.
- For example, cloud hosting control plane functionalities and network intelligent hypervisor to configure network routing node



Example cloud hosting control plane functionalities and network intelligent hypervisor to configure network routing node



Summary and Conclusions

Summary of the enablers (1/2)



Enabler	Background	Benefits	Implications
Architectural means and protocols	6G data-driven architecture will require architectural support that enables communication for AI	Can help define the inter-layer APIs and the protocols used to connect the layers of an E2E system design	Define internal and external APIs that realize the inter-layer interaction
MLOps	AI execution environments will be everywhere (e.g., UE, RAN and Core) and require tools for managing the lifecycle of these AI models	Improve operation, management, and maintenance of the E2E system design	Privacy-aware data collection, AI management.
AlaaS	AlaaS is a framework that offers a wide range of AI services as well as personalized inference capabilities to the AI service itself	Improve M&O and FCAPS of the network as well as impact on design of the E2E system	Impact on the E2E system design, AlaaS needs DataOps, MLOps and protocols
DataOps	Data shall be delivered, pre-processed, and stored where and when required. This imposes requirements on a flexible data ingestion architecture.	Efficiently collect and process data, as well as provide inferences to the data consumers within the E2E system design	Impact on the RAN and CN architecture; functions, protocols and interfaces may be needed.

Enabler	Background	Benefits	Implications
Optimized network function composition	Modular design minimizes the dependencies between different modules while the relevance of the NF functionalities within a module is maximized	Increased flexibility, optimized signaling, and efficient resource usage	Impact on the CN NF design in 6G since the design need to be different from 5G
Streamlined network function interfaces & interaction	The network modules and their interfaces need to support the coexistence of these use cases as well as the related services.	Extend the support for new and existing use cases as they could be optimized based on the NF (or Network module) placement choices (e.g., centralized and distributed cloud deployments).	Impact on NF design and 5G procedures. Need for new interfaces and interaction
Flexible feature development and run-time scalability with modular network functionality	Exploring the possible enhancements to the E2E modularization (e.g., network slicing in 5G) to optimize the functionality	Enhanced network slicing & performance, flexibility via modularization, customization of E2E functionality	E2E impacts as the design and placement of network modules through the cloud continuum (e.g., cloud, edge, access, extreme edge etc.) would be revisited.
Network autonomy & Multi-X orchestration	In 5G, network slicing was a key enabler to facilitate the co-existence of various use cases with demanding and often conflicting requirements. The management and orchestration are built upon open loop slice configurations and semi-static parameters from SLAs which often result in low resource utilization	Improved data-based slice management. With a more autonomic and closed-loop based slice orchestration mechanisms it will be possible to address the orchestration of the network services including the extreme-edge domain, which is highly dynamic, heterogeneous, and volatile.	E2E impacts as the NF placement decisions through cloud continuum would be optimized with a higher time granularity based on the network dynamics. It requires closed-loop control and more flexible orchestration mechanisms as well as an enhanced exposure process. It will require also to define a comprehensive information model capturing the peculiarities of those devices in extreme-edge domain.
Network migration	To perform this transition as efficiently as possible 6G should take this into consideration from the beginning: how to migrate from 5G to 6G	Critical, will determine how the E2E system will look like	It has fundamental impact to the 6G RAN & CN as it will outline the evolution path from 5G to 6G

Enabler	Background	Benefits	Implications
Network of networks	Integration of multiple subnetworks, including terrestrial and non-terrestrial networks in order to create a seamless and ubiquitous communication system	Improved coverage, reduced complexity, increased reliability and more efficient management of network resources	New UE roles and responsibilities in a subnetwork, communication between non-terrestrial nodes, trust of diverse network nodes, communication and computation resource management
Multi-connectivity	Multi-connectivity enables multiple frequency ranges by different physically separated nodes, the aggregation of different radio access technologies, carriers, and access networks	Robustness and reliability, increased throughput	Depending on the solution, new interfaces and protocols between nodes may be needed, which may lead to an increased complexity in coordinating different NW nodes. New mechanisms and procedures for integration of multiple RANs should be defined
E2E context awareness management	Managing a network so that it can adapt dynamically to the context	Mission-critical operations to reduce the network overhead and to allocate edge resources flexibly, ultimately improving the system performance by allowing multiple edge allocations and RAN slices.	Different network components e.g., RAN/CN, transport, applications, should become aware of the context and need to interact, implying the need for signalling and synchronisation. Effective resource allocation and orchestration mechanisms that operate even when incomplete or partial context awareness is available should be designed

AI enablers for data-driven architecture

Network modularisation

Architectural enablers for new access and flexible topologies

Summary of the enablers (2/2)



Enabler	Background	Benefits	Implications
Exposure and data management	Functions to process the data collected and how to expose the (managed) data to external an internal usage	Expose data that may enable new 6G services	Impact mainly on the CN architecture and the Network-centric application layer; new functions, protocols and interfaces may be needed
Protocols, signalling and procedures	Discovery of compute nodes and impact of new sensing services on RAN interfaces and functionality	Critical to implement the Beyond Comm. Functionalities (BCFs)	New radio measurements needed; protocols needed to collect data to the data management.
Application- and Device-driven optimisation for Beyond Communication Services	Defining the requirements associated to applications using JCAS, Digital Twinning, etc	Improved QoS/QoE through efficient placement of BCF/BCS data/inference consumers (application, devices) within the E2E system design	Enhanced orchestration mechanisms across the continuum; efficient network – application component communication and service exposure.
Enhancing Joint Communication and Sensing Capabilities	Several 6G use cases require extreme localization performance, such as highly precise SLAM. Furthermore, massive distributed JCAS can raise significant limitations for classical technologies; quantum technologies is thus needed.	Accurate indoor mapping in challenging scenarios with the help of sensing information from COTS devices. Overcoming limitations for classical technologies on massive communication overhead and computational complexity	Impacts the sensing design and data collection from COTS devices for SLAM; Hybrid classical-quantum network, where quantum virtual machines will hold entanglements and qubits for its usage in the network

Enabler	Background	Benefits	Implications
Integration and orchestration of computing continuum resources into the 6G architecture	Future 6G networks will consider computing capabilities across the full network, from the extreme edge (including the UE) to Telco grade clouds (CC, Compute Continuum)	Better management of the resources and services in the CC	Impacts the extension of the CC, where strong emphasis is given on the extreme-edge integration, management and usability
Multi-domain/Multi-cloud federation	Different Telco-Cloud Providers are offering Compute, Storage and Network resources as a service on different platforms complicates management	Unification of existing domain specific orchestration frameworks into an e2e federated architecture	It has fundamental impact in addressing the exiting challenges in multi domain federation as it will define important design, integration and orchestration principles for federation of services or/and resources.
Network modules placement in the resource continuum	The function (module) placement in the CC is challenging considering the heterogeneity and volatility of the Edge and Extrema Edge resources	Improve flexibility and critical services of the network by real-time function placement	Definition of APIs to expose capabilities of heterogeneous computing resources in the CC
Cloud Transformation in 6G-quantum architecture	The 6G edge supporting the softwarized network continuum will imply the explosion of network control traffic. Quantum technologies and a 6G-quantum network architecture can improve the optimal use of resources.	By using lossy encoding of data in quantum bits and adapting the cloud with algorithms to process them, it is possible to reduce the load of data mining procedures.	Will impact the reduction of traffic load in the CC by integrating quantum technologies

Network beyond communications

Virtualisation and cloud continuum transformation



References an Acronyms

References



- [ERI-MLOps] Ericsson Research Blogpost, “Decentralized learning and intelligent automation: the key to zero-touch networks?”, March 2022.
- [ERI-Data] Ericsson Technology Review, “Data Ingestion Architecture for Telecom Applications,” March 2021.
- [ERI-ZTN] Ericsson Technology Review, “Creating Autonomous Networks with Intent-Based Closed-Loops”, April 2022.
- [36.888] 3GPP TR 36.888, “Study on provision of low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE (Release 12)”, June 2013.
- [37.340] 3GPP TS 37.340, “Multi-connectivity (Release 17)”, April 2023
- [38.821] 3GPP TR 38.821 “Technical Specification Group Radio Access Network; Solutions for NR to support non-terrestrial networks (NTN)”, V16.1.0 (2021-05)
- [FU98] G. D. Forney and G. Ungerboeck, “Modulation and coding for linear Gaussian channels”, IEEE Transactions on Information Theory, vol. 44, no. 6, pp. 2384-2415, October 1998.
- [Maz75] J. E. Mazo, “Faster-than-Nyquist signaling”, Bell System Technical Journal, vol. 54, no. 8, pp. 1451-1462, October 1975.
- [Moe21] K. Moerman. Joint Communication and Sensing in 6G Networks. <https://www.nxp.com/company/blog/joint-communication-and-sensing-in-6g-networks:BL-SENSING-IN-6G-NETWORKS>, Nov. 5, 2021
- [TAZ+13] A. Tzanakaki, M. P. Anastasopoulos, G. S. Zervas, B. R. Rofoee, R. Nejabati and D. Simeonidou, “Virtualization of heterogeneous wireless-optical network and IT infrastructures in support of cloud and mobile cloud services”. IEEE Communications Magazine, vol. 51, no. 8, pp. 155-161, August 2013.
- [BA+21] Bassoli, R. et al. (2021). Quantum Communication Networks. Foundations in Signal Processing, Communications and Networking, vol. 23. Springer, Cham.
- [GKM+22] M. Gramaglia, M. Kajo, C. Mannweiler, O. Bulakci and Q. Wei, “A unified service-based capability exposure framework for closed-loop network automation”, *Transactions on Emerging Telecommunications Technologies*, vol. 33, no 11, July 2022.
- [HEX22-D62] Hexa-X Deliverable D6.2: https://hexa-x.eu/wp-content/uploads/2022/05/Hexa-X_D6.2_V1.1.pdf
- [SCT+22] A. M. Sanchez, A. S. Charismiadis, D. Tsolkas, D.A. Guillen, and J.G. Rodrigo, “Offering the 3GPP Common API Framework as Microservice to Vertical Industries”, *2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)* (pp. 363-368). IEEE, June 2022.
- [ZSM002] ETSI GS ZSM 002, “Zero-touch network and Service Management; Reference Architecture”, August 2019.
- [OR22] Ordóñez-Lucena, J., & Dsouza, F. (2022, August). Pathways towards network-as-a-service: the CAMARA project. In *Proceedings of the ACM SIGCOMM Workshop on Network-Application Integration* (pp. 53-59).
- [HEX22-D13] Hexa-X Deliverable D1.3: https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-X_D1.3.pdf
- [HEX22-D14] Hexa-X Deliverable D1.4
- [HEX22-D62] Hexa-X Deliverable D6.2: https://hexa-x.eu/wp-content/uploads/2022/05/Hexa-X_D6.2_V1.1.pdf
- [Guer16] R. Guerzoni, et al, “Multi-domain Orchestration and Management of Software Defined Infrastructures: a Bottom-Up Approach”, 2016, Available: http://real.mtak.hu/41606/1/5GEx_bottom_up_approach_v1_3_u.pdf
- [WSL+21] H. Wymeersch, D. Shrestha, C. M. de Lima, V. Yajnanarayana et al., “Integration of communication and sensing in 6g: a joint industrial and academic perspective,” 2021.

References



- [HEXA-I-D5.2] Hexa-X Deliverable D5.2, “Analysis of 6G architectural enablers applicability and initial technological solutions”, October, 2022.
- [HEXA-I-D5.3] Hexa-X Deliverable D5.3, “Final 6G architectural enablers and technological solutions”, April 2023
- [HEX23-D21] Hexa-X-II Deliverable D2.1, “Draft foundation for 6G system design”, to be published June 2023
- [JSAC-1] Joint Communications & Sensing, Common Radio-Communications and sensor technology, VDE ITG [Available]: <https://www.vde.com/resource/blob/2108086/e9ead7b7c95eacb5cb7f2eb88103eaa9/joint-communication---sensing-position-paper-pdf-file-data.pdf>

Acronyms



- 6GC: 6G Core
- AGV: Automated Guided Vehicle,
- AI: Artificial Intelligence
- AlaaS: AI as a Service
- AMF: Access and mobility management function
- AUSF: Authentication server function
- API: Application Programming Interface
- BSS: Business Support Solutions
- CA: Carrier Aggregation
- CaaS: Compute-as-a-Service
- CN: Core Network
- CP: Control Plane
- CP-G: Control Plane Gateway
- CU: Centralized Unit
- D2D: Device-to-Device
- DataOps: Data Operations
- DC: Dual Connectivity
- DL: Downlink
- E2E: End to end
- EE: Extreme Edge
- EN-DC: EUTRA-NR Dual Connectivity
- ETSI: European Telecommunications Standards Institute
- FA: Federation Agent
- GIS: Geographic Information System
- GPRS: General Packet Radio Service
- GTP: GPRS Tunnelling Protocol
- ISL: Inter-Satellite Link
- KPI: Key Performance Indicator
- KVI: Key Value Indicators
- LoS: Line of Sight
- MA: Management Agent
- MaaS: Mobility as a Service
- MC: Multi-Connectivity
- M&O: Management and Orchestration
- MEC: Multi-access Edge Computing
- mMIMO: Massive Multiple-Input Multiple-Output
- MCG: Master Cell Group
- ML: Machine Learning
- MLOps: Machine Learning Operations
- MN: Management Node
- MR-DC: Multi-Radio Dual Connectivity
- NB: NodeB
- NF: Network Function
- NRF: Network repository function
- NS: Network Slice
- NSS: Network Slice Subnet
- NTN: Non-Terrestrial Network
- NWDAF: Network Data Analytics Function
- OSS: Operational support solutions
- RaaS: Robotic-as-a-Service,
- RAN: Radio Access Network
- RRC: Radio Resource Control
- RRM: Radio Resource Management
- SCG: Secondary Cell Group
- SDN: Software Defined Network
- SBA: Service Based Architecture
- SDO: Standards Development Organizations
- SMF: Session management function
- SL: Service Link
- SoA: State-of-Art
- TL: Transport Layer
- TN: Terrestrial Network
- TN: Transport Network
- ToD: Teleoperated Driving
- UE: User Equipment
- UL: Uplink
- UP: User Plane
- UPF : User Plane function
- UP-G: User Plane Gateway
- VR: Virtual Reality
- ZSM: Zero touch network and service management
- ZT: Zero Touch



HEXA-X-II.EU //   



Co-funded by
the European Union

6GSNS

Hexa-X-II project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101095759.