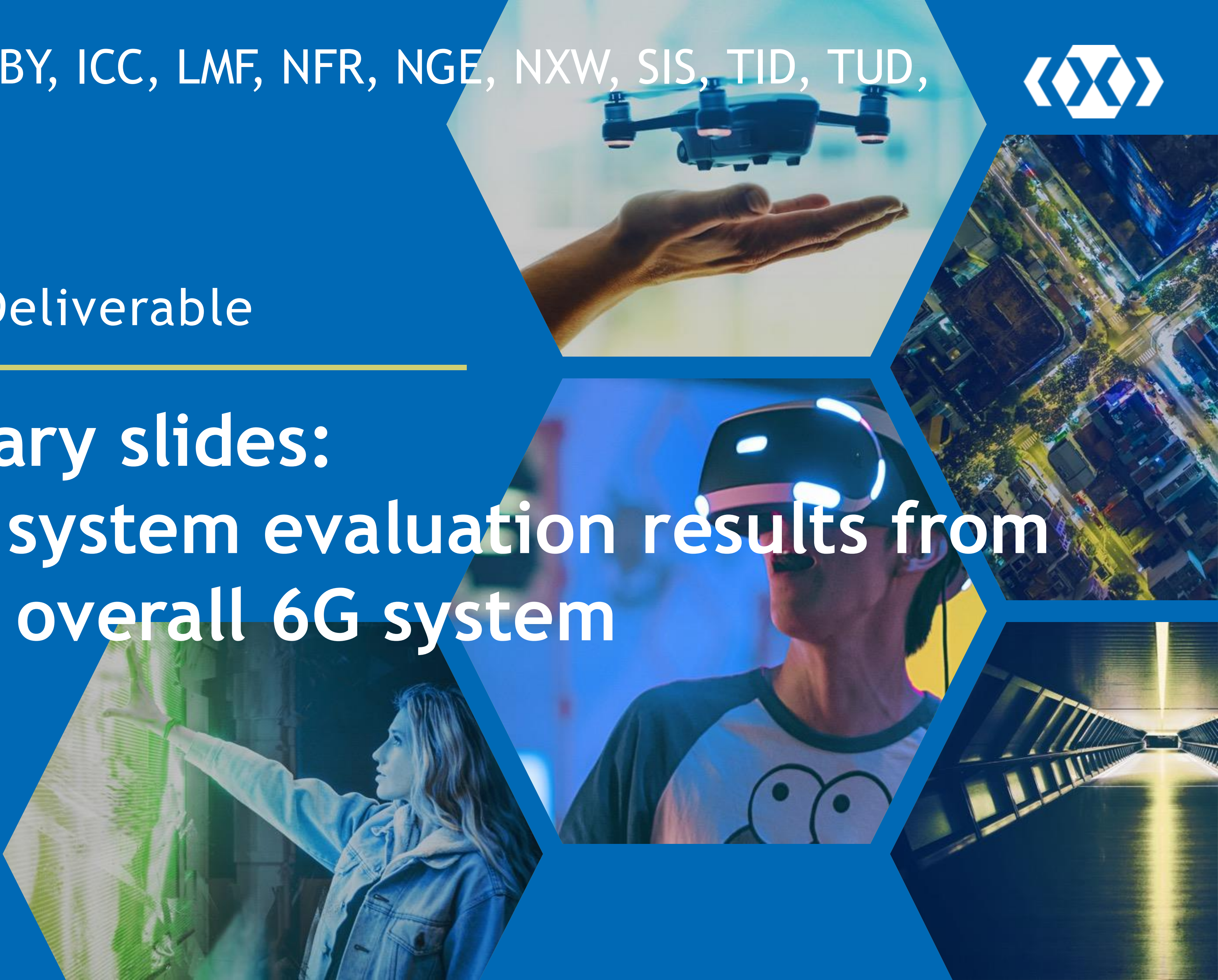
Partners: BI, CTT, EBY, ICC, LMF, NFR, NGE, NXW, SIS, TID, TUD, VTT, WIN, EAB

Hexa-X-II D2.4 Deliverable

D2.4 summary slides: End-to-end system evaluation results from the interim overall 6G system



Hexa-X-II

hexa-x-ii.eu

2024-09-30

Table of content

Chapter 1: Introduction

Chapter 2: Introduction to E2E System

Chapter 3: E2E system-level evaluation results

Chapter 4: Evaluation of the security, privacy and system-level resilience

Chapter 5: Conclusions

Chapter 1 Introduction

Introduction

This document aims to become a transitional outcome within the Hexa-X-II lifetime:

- On the one hand, this deliverable aims to describe and illustrate the set of results and outcomes originated from all the work presented in the previous deliverables.
- On the other hand, this deliverable will allow to trigger the last phase of the 6G E2E System blueprint study and the System-PoC implementations work that will lead towards the future publication of the last two WP2 deliverables (i.e., D2.5 and D2.6), which will present the last set of outcomes produced by the Hexa-X-II project.

Objectives

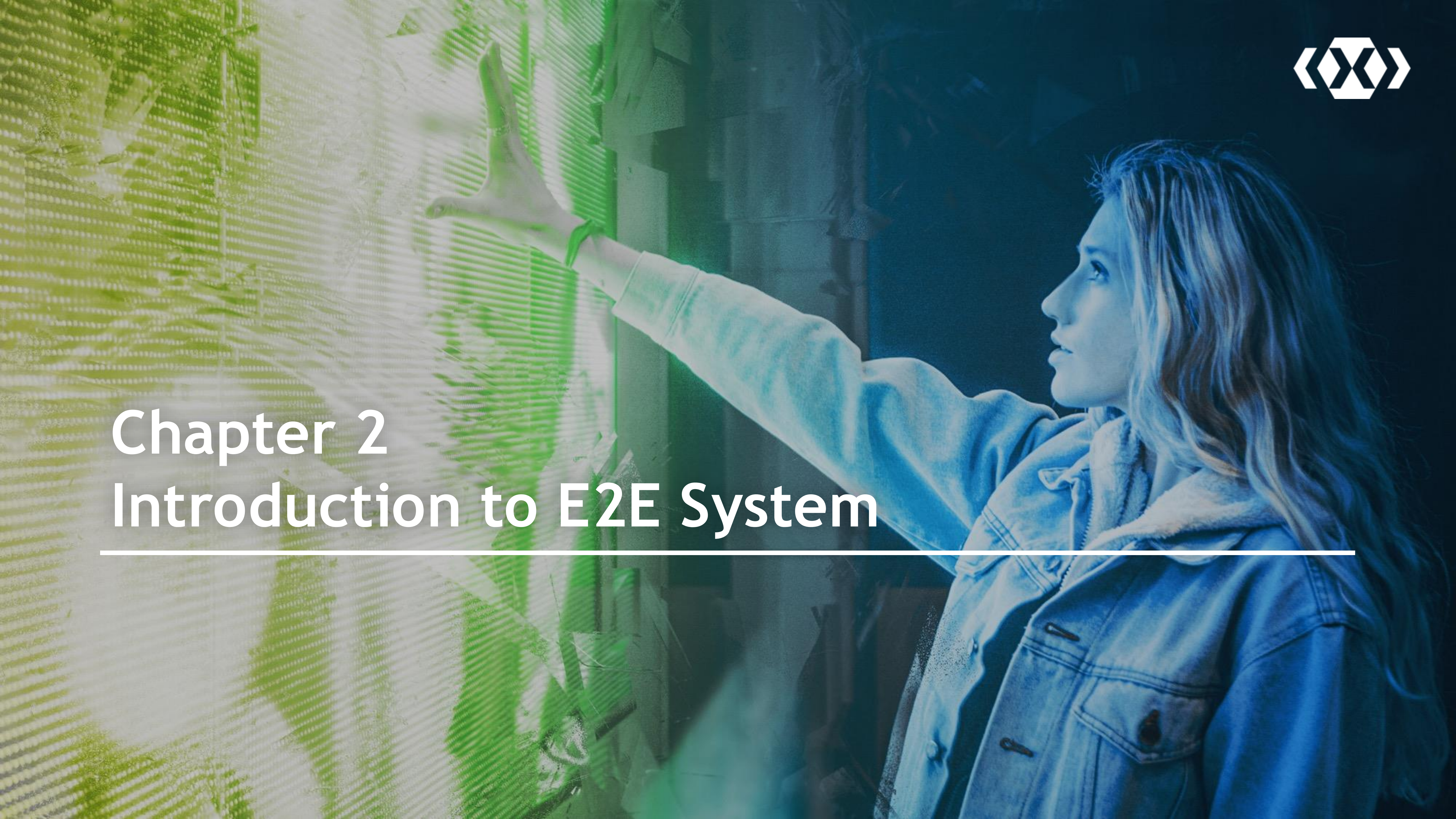
- Objective 1: To introduce the current 6G E2E system design status with an overview of its elements and the different layers composing it, together with those aspects that should be evaluated and validated.

- **Objective 2:** To provide the next round of experimental outcomes obtained from an E2E system point of view. This objective is divided in two sub-objectives:
  - Sub-Objective 2.1: To present the experimental results related to the System-PoC #B and its components associated. In this deliverable a clear evolution is presented since the outcomes illustrated in D2.3.
  - Sub-Objective 2.2: To present the results achieved by the set of security-related enablers introduced in D2.3 with the focus of enforcing security on 6G E2E systems.

- **Objective 3:** To conclude how the different experiments and their results allow to validate the different aspects presented in objective 1.

Deliverable structure

- **Chapter 2** provides a brief introduction and description of the latest E2E System design and the aspects under evaluation.
- **Chapter 3** presents the first set of the E2E system-level evaluation results with special focus on the System-PoC B; the designed architecture with its components and the implemented scenarios used to obtain the experimental outcomes.
- **Chapter 4** describes the obtained experimental outcomes from the security enablers presented in Hexa-X-II deliverable D2.3 [HEX224-D23], bringing topics such as anomaly detection, Joint Communication and Sensing (JCAS) threat mitigation, trust aspects and the use of Artificial Intelligence (AI).
- **Chapter 5** presents the conclusions of this document, with special focus on the 6G E2E System evaluation aspects based on the presented results.

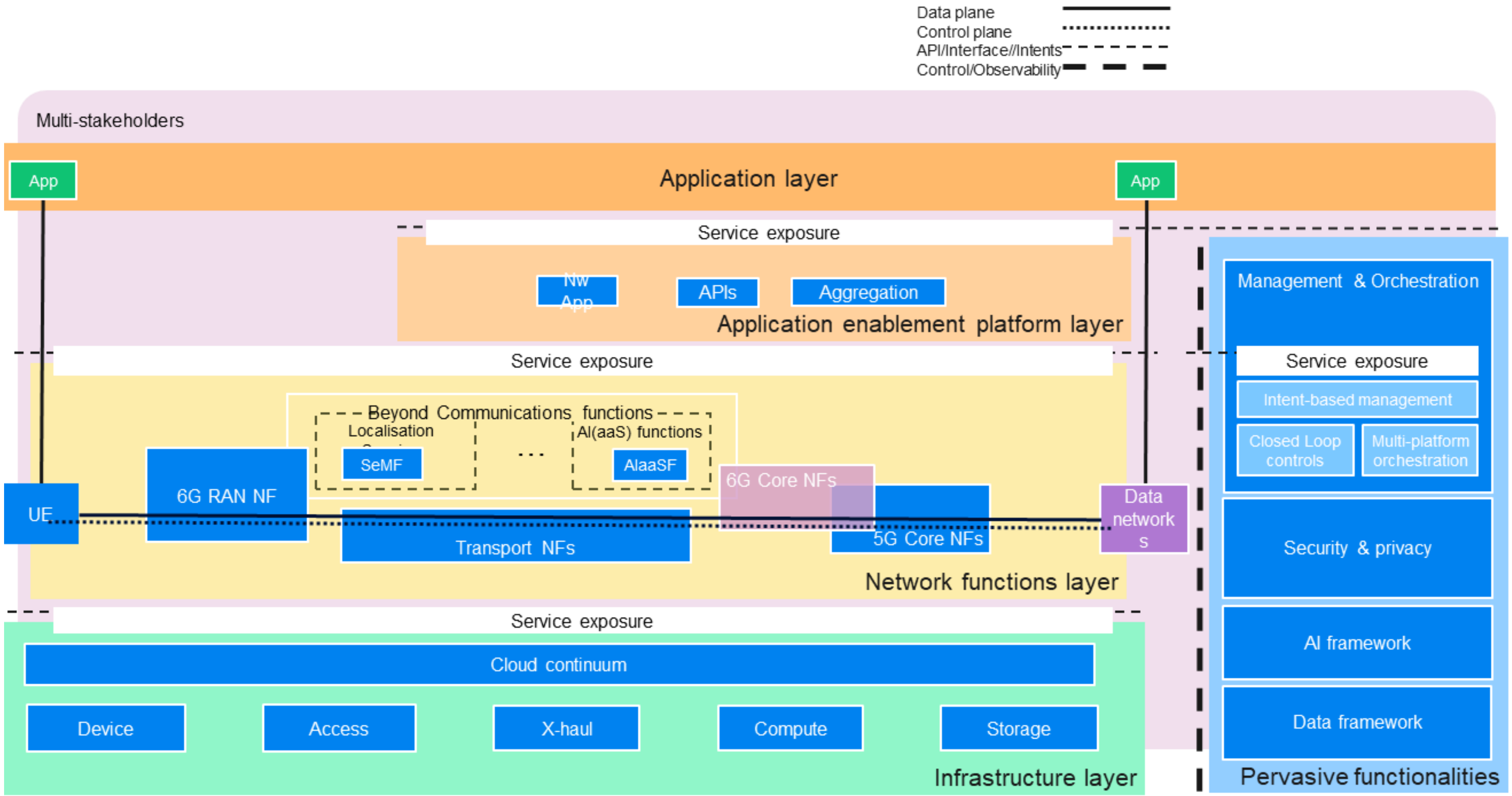


Chapter 2
Introduction to E2E System

E2E System design overview

The iterative system blueprint design

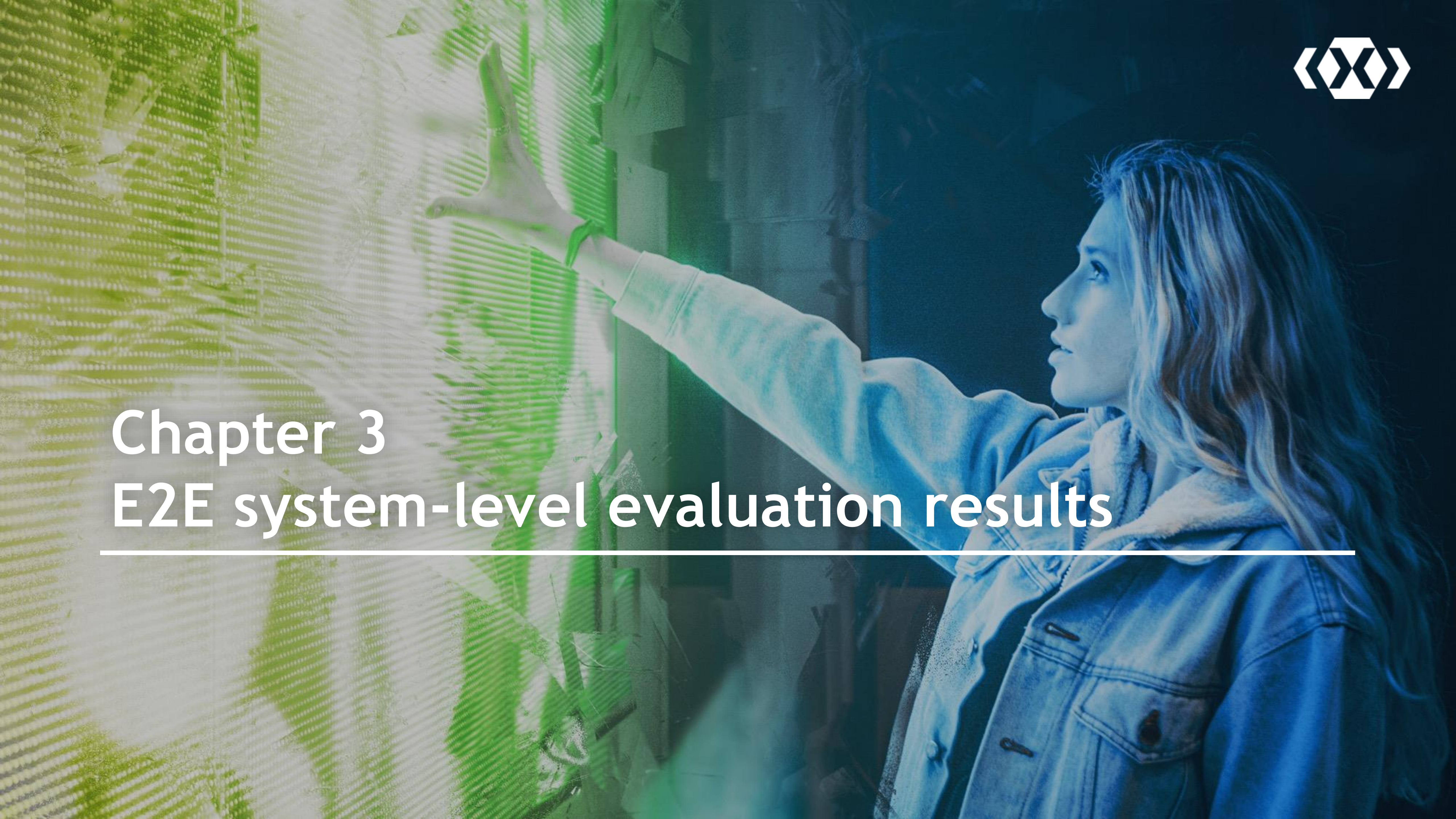
- Maps dependencies between enablers, layers, and pervasive functionalities.
- Evaluates through system-level proof of concepts or simulations to estimate KPIs and KVIs.
- Iteratively improves system design based on evaluation results



E2E System design aspects under evaluation

Design principle	Aspects under evaluation
Principle 1: Support and exposure of 6G services and capabilities	• Compute-as-a-service, AI-as –a-Service exposure across domains and layers. • Intent-based management with intent-based APIs for 6G customers. • Management capabilities exposure framework for Hexa-X-II capability interoperability
Principle 2: Full automation and optimization	• Pervasive infrastructure & resource orchestration across the compute continuum. • Comprehensive data fusion, monitoring, and telemetry in data framework. • Multi-platform orchestration box for synergetic orchestration mechanisms. • AI/ML-based algorithms for closed loop control and optimization of real-time services
Principle 3: Flexibility to different network scenarios	• Multi-agent and multi-cluster orchestration for flexible service deployment. • Federated and decentralized orchestration for dynamic connections to 3rd party networks. • Network programmability for flexible topologies and backhaul connectivity. • Application awareness and adaptive QoS/QoE.
Principle 4: Scalability	• Pervasive service management and orchestration system for scaling based on traffic and mobility needs. • Decentralized orchestration components for both small and large-scale deployments. • Network programmability enabler for scalability of optimized transport network functions
Principle 5: Resilience and availability	• Reliability and availability of the system under different conditions. • Data analysis, AI integration, and coordination mechanisms in ensuring resilience
Principle 6: Persistent security and privacy	• Integrate resilience functionalities to address current as well as future threats. • Inherently support the preservation of privacy. • Allow different levels of anonymity for future services.
Principle 7: Internal interfaces are cloud optimized	• Could native APIs for services (including management & orchestration services) exposure and consumption.
Principle 10: Minimize environmental footprint and enabling sustainable use cases	• Implementation of a pervasive data and analysis framework. • Infrastructure layer in the 6G E2E system with optimized energy consumption and costs for enhanced sustainability and operational efficiency of the use cases.

Note: Principles 8 (Separation of concern of network functions) and 9 (network simplification in comparison to previous generations) are not evaluated in this deliverable.



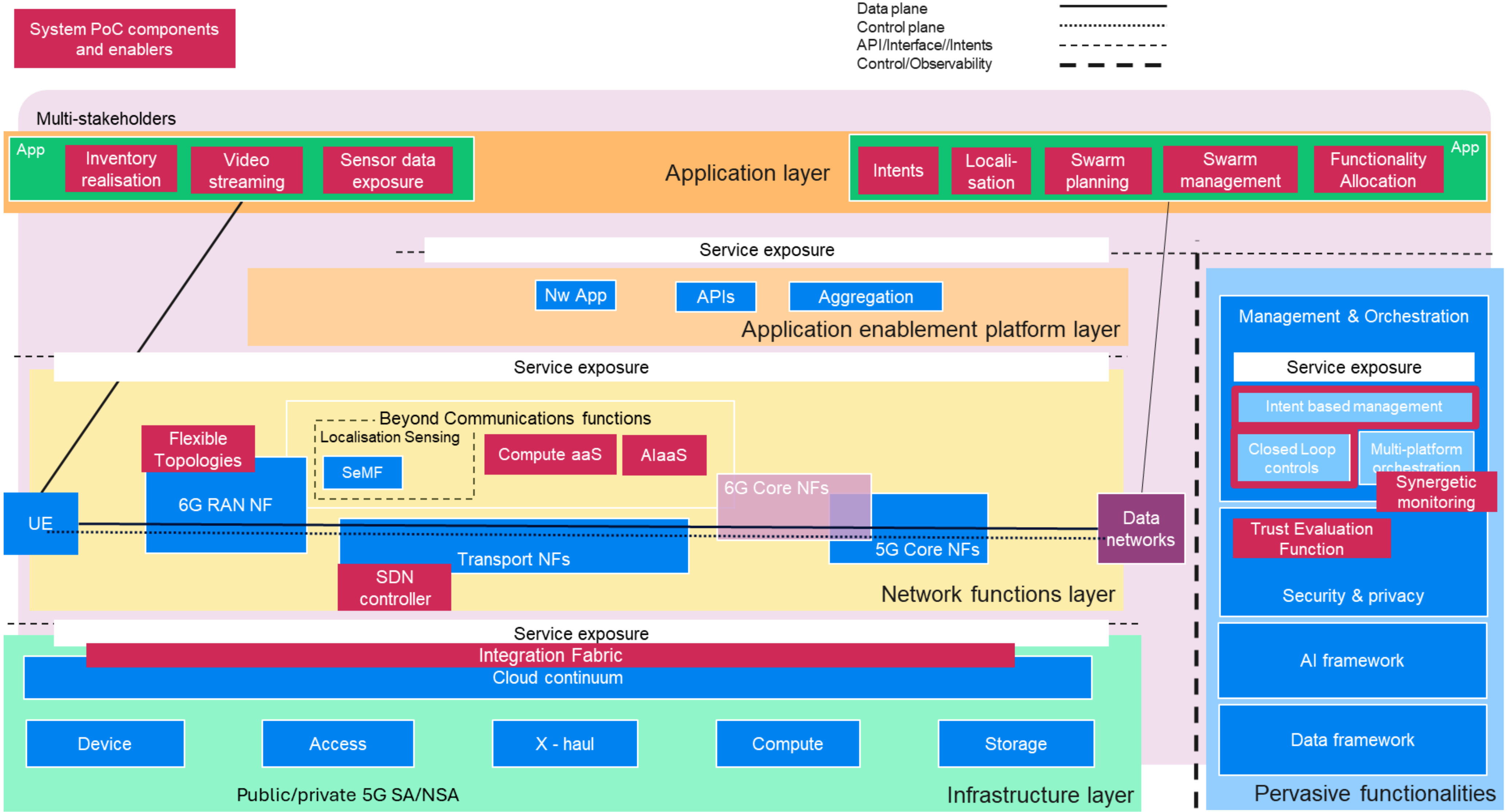
Chapter 3
E2E system-level evaluation results

Overview of the E2E system evaluation and validation activities

- The **first iteration (System-PoC #A):**
  - Detailed description and results are reported in D2.2 and D2.3
  - Focused on smart network management aspects for demonstrating management mechanisms.
- The **second iteration (System-PoC #B)**
  - Designed to build upon System-PoC #A by incorporating new enablers and assuming that not all of the network and cloud domain resources, application components and devices are part of the same domain and/or geographical location.
  - Aims to to orchestrate all involved components in a coherent and synchronised way.
  - Allows to broaden the sustainability aspects already considered in System-PoC #A, e.g., social sustainability, environment sustainability, etc.
- The **third iteration (System-PoC #C)**
  - Focuses on adding radio and device aspects upon System-Poc #B.
  - To be reported in future deliverables (D2.5 and D2.6).

System PoC e2E architecture and alignment with the 6G system blueprint

A mapping of the System-PoC's functional and performance aspects (red boxes) to the 6G system's blueprint.



Components of System-PoC #B

Application and M&O features

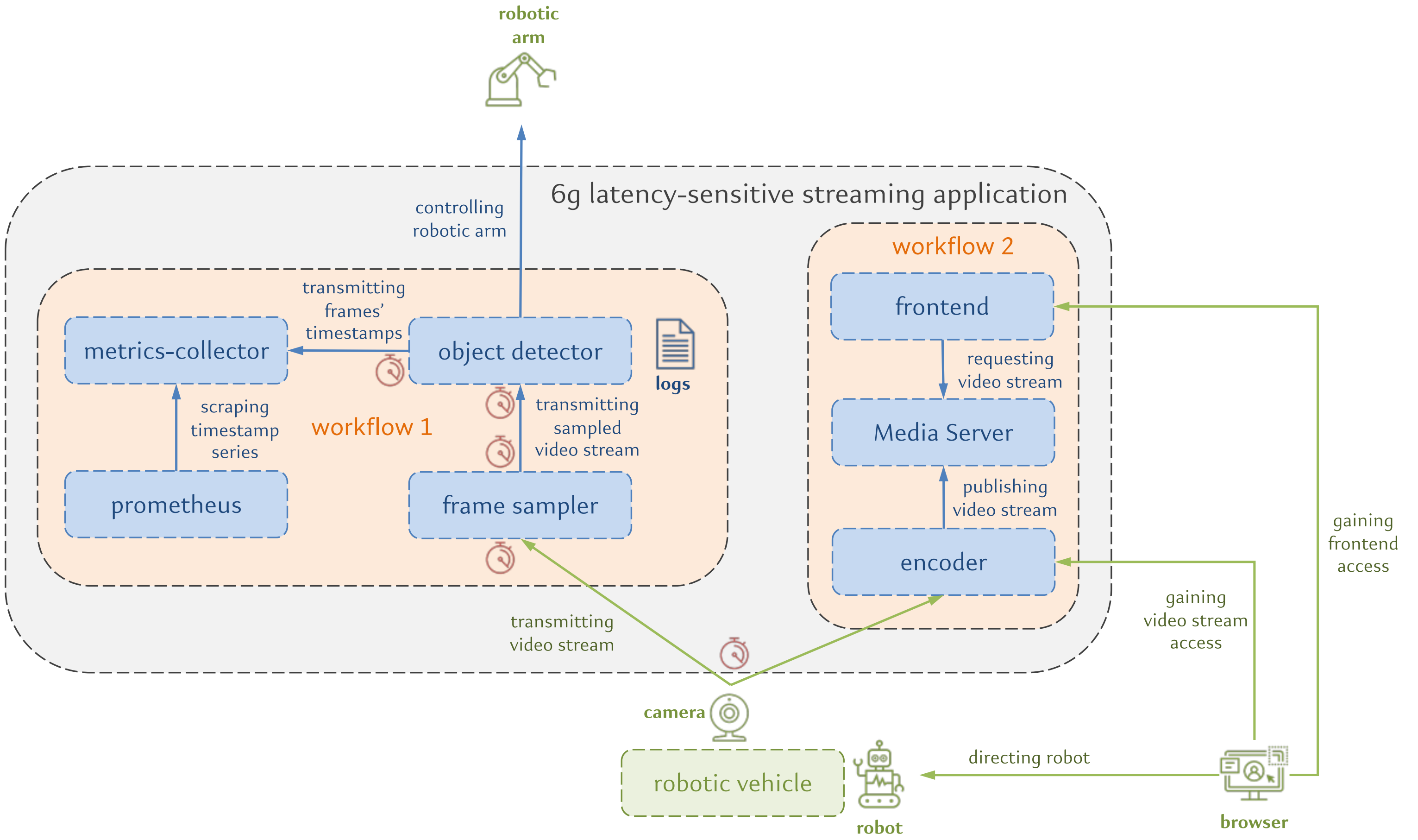
Application components of the Automated Inventory Management solution:

- **Inventory realisation**: Identifies and tracking of the location (e.g., warehouse) items (including computer vision modules).
- **Video streaming**: Implements the transfer of the AMR/UAV-oriented data streams towards further processing.
- **Sensor data exposure**: Enables the AMR/UAV/infrastructure sensor information to be accessible by the respective data consumers.
- **Intents**: Enables the end-user to insert high-level requests/commands towards the system, related to the inventory operations' performance, duration, resource requirements, etc.
- **Localisation**: Provides location information to the respective consumer modules (e.g., swarm planning and management).
- **Swarm planning/deployment**: Implements the functionality related to the AMR/UAV resources reservation to perform the inventory operations.
- **Swarm management**: Monitors the swarm during runtime and identificaties and handles the events that arise (e.g., battery depletion, hardware malfunctions, etc.)
- **Functionality allocation (FA):** Optimizes the placement of a) the inventory management; b) workloads requiring considerable computational resources. Reported in D2.2, D2.3.
- **Trust Evaluation Function: A**nalyses data (real-time and historical) coming from the compute nodes of the system and assessing the trust indices of each node that are used by the FA.
- **Synergetic Monitoring**: Exposes the infrastructure, device, and network monitoring metrics of multiple domains at a common location towards the joint optimization of resources.

Application and M&O features

6G latency sensitive application consisting of two workflows:

1) A streaming pipeline for live streaming the footage from the robot to the frontend, and
2) An object detection which is sampling the frames and detecting possible danger situations in the manufacturing facility.



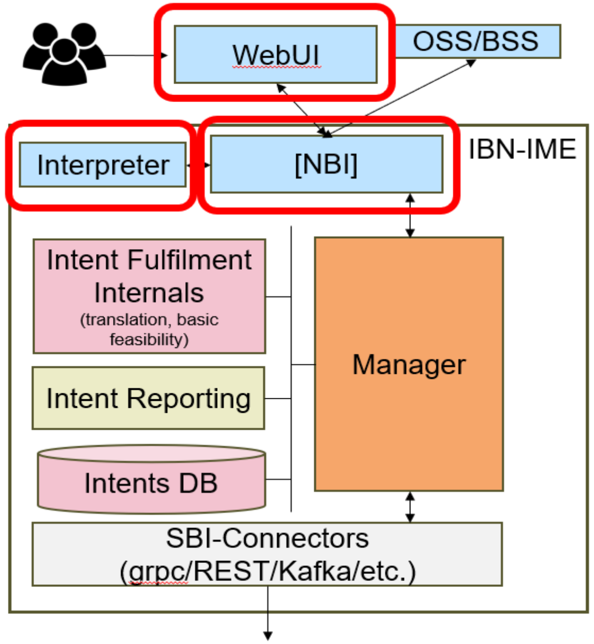
Intent-based Network solution

- IBN module to manage and process intents and its associated data objects (intents and intent reports).
- Use of Natural Language Processing (NLP) to process human requests and generate intents based on 3GPP data models.
- Interaction with different OSS/OSC through multiple South Bound Interfaces (SBI). → E.G., REC-EXEC
- Use of reporting functions to deliver a complete intent-based functionality.

*IBN-IME: Intent-based Management - Intent Management Entity

"create a cobot application on warehouse 24 with zero downtime"

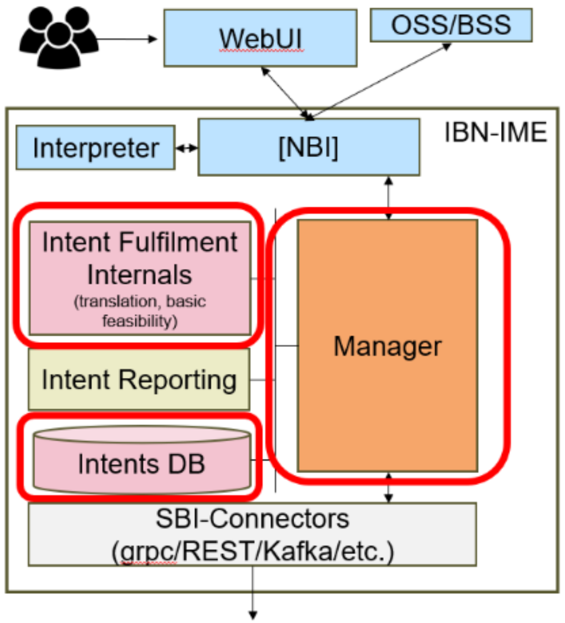
Intent reception and interpretation tasks.



WebUI
OSS/BSS
Interpreter
[NBI]
IBN-IME
Intent Fulfilment Internals
(translation, basic feasibility)
Intent Reporting
Intents DB
Manager
SBI-Connectors (grpc/REST/Kafka/etc.)

- Identified Operation: create
- Identified Resource: cobot application
- Identified Location: warehouse 24
- Identified Recursivity: True (since it requested zero downtime)

Intent creation, feasibility, translation and storage tasks.



WebUI
OSS/BSS
Interpreter
[NBI]
IBN-IME
Intent Fulfilment Internals
(translation, basic feasibility)
Intent Reporting
Intents DB
Manager
SBI-Connectors (grpc/REST/Kafka/etc.)

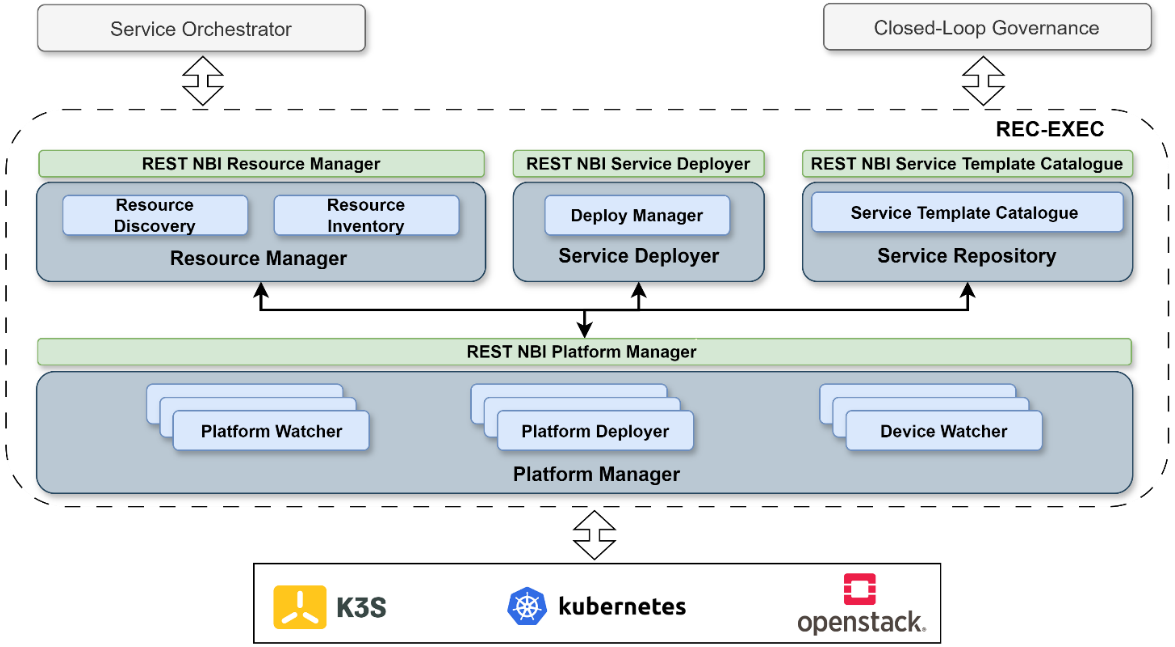
Role in the PoC

- Offer an intent-based solution to avoid the need of having technical knowledge at the client side, leaving all the resources and service management to the system.

(Service and) Resource orchestration

Integration of two dedicated orchestrators for the **services** and the **resources**

- **Resource orchestrator (REC-EXEC) →** implementation of the *Multi-cluster resource manager* enabler
  - ○ **Resource Manager.** Stores information related to the platforms and the devices.
  - ○ **Service Deployer.** In charge of application deployment (vertical applications, closed-loop functions, etc.) on the target cloud platform.
  - ○ **Service Repository.** Stores platform-specific orchestration templates e.g., Helm charts for K8s
  - ○ **Platform Manager.** Encompasses a set of technology-specific driver to interact with the different target cloud platform



- **Service Orchestrator** is a blackbox

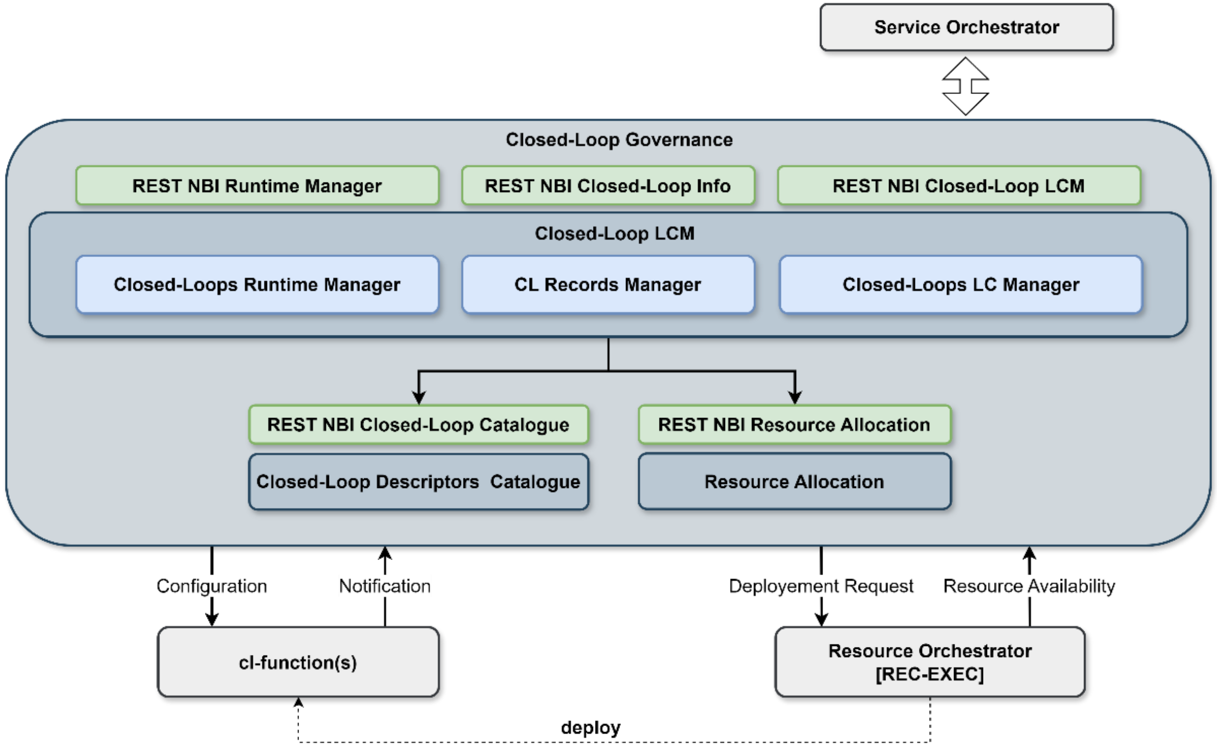
Role in the PoC

- Provision a Video Streaming application on a target cobot through the dedicated platform.
- Request the cobot platform to start the target cobot's patrolling.
- Provision Closed-Loop functions at the edge cloud (a Kubernetes cluster) to enable service automation.

Closed-loop automation

Closed-loop (CL) Governance internal structure

- **CL Descriptors Catalogue**. Repository storing CL and CL functions descriptors.
- **Resource Allocation.** In charge of computing CL functions' placement by exploiting information retrieved from the REC-EXEC
- **CL LCM**. Responsible for the lifecycle management of the CLs. 3 main sub-modules:
  - **CL LC Manager.** Selects proper CL and CL functions' descriptors given the service requirements, provisions and decommissions CLs
  - **CL Records Manager.** Tracks the existing instances of CL.



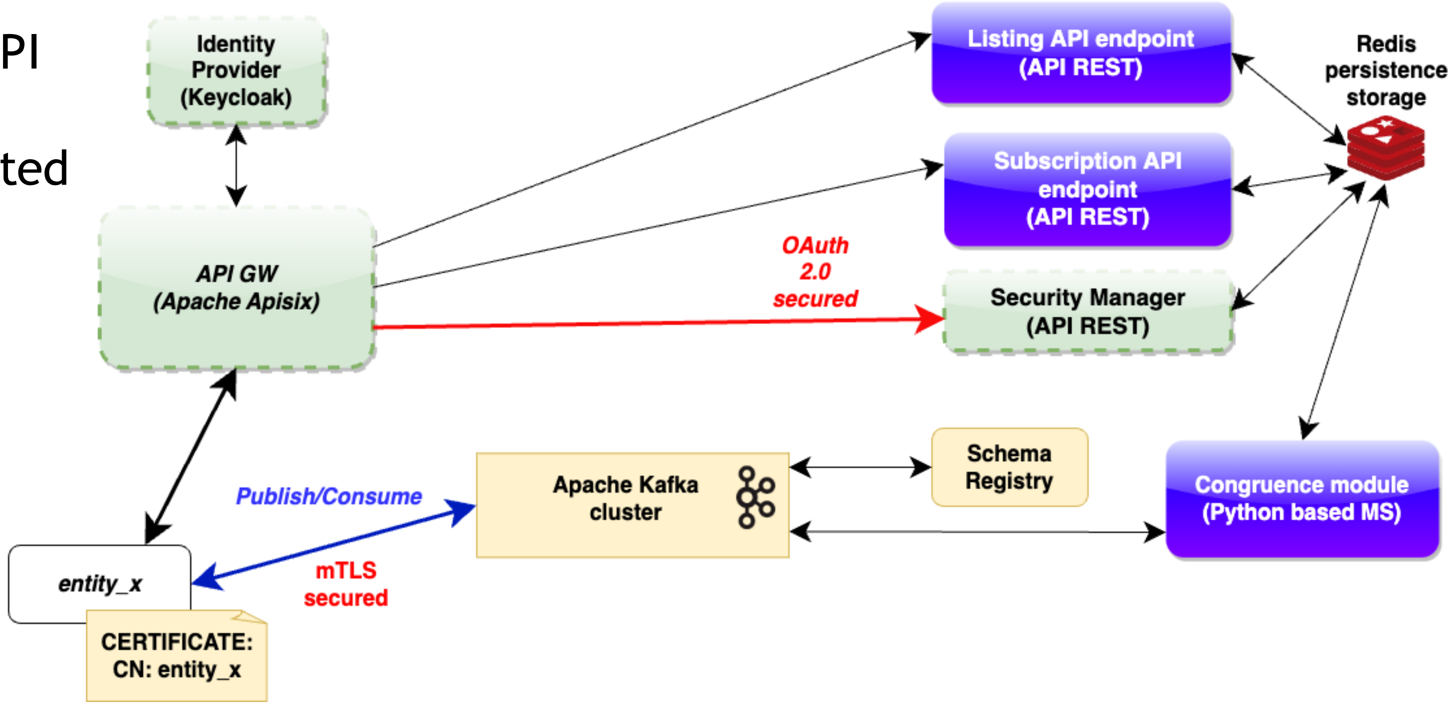
  - **CL Runtime Manager.** Performs runtime operation of existing CL instances, e.g., configuration, start/stop, etc.

Role in the PoC

- In charge of the management of a specific CL to enable the zero-touch automation for the cobot-based surveillance service.

Management Capabilities Exposure Framework (MCEF)

- MCEF provides a structured communication channel that exposes management capabilities within the E2E architecture to third-party entities.
- Based on an Apache Kafka cluster and REST API endpoints.
- Employs mutual TLS (mTLS) to ensure encrypted communication.
- Access is controlled via Access Control Lists (ACLs) based on the entity's credentials.
- Offers standardized REST APIs (OpenAPI 3).



Identity Provider (Keycloak)
API GW (Apache Apisix)
Listing API endpoint (API REST)
Subscription API endpoint (API REST)
Security Manager (API REST)
Redis persistence storage
OAuth 2.0 secured
Publish/Consume
mTLS secured
Apache Kafka cluster
Schema Registry
Congruence module (Python based MS)
entity_x
CERTIFICATE: CN: entity_x

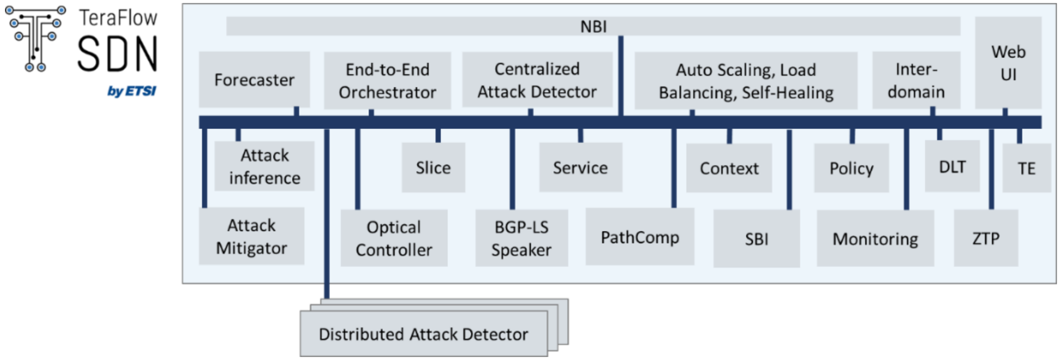
Role in the PoC

- **Capability Exposure:** Securely exposes management capabilities to both internal components and third-party entities.
- **Interoperability and Security:** Enables interoperable communication while enforcing robust security, ensuring that only authorized entities can access and interact with the system.

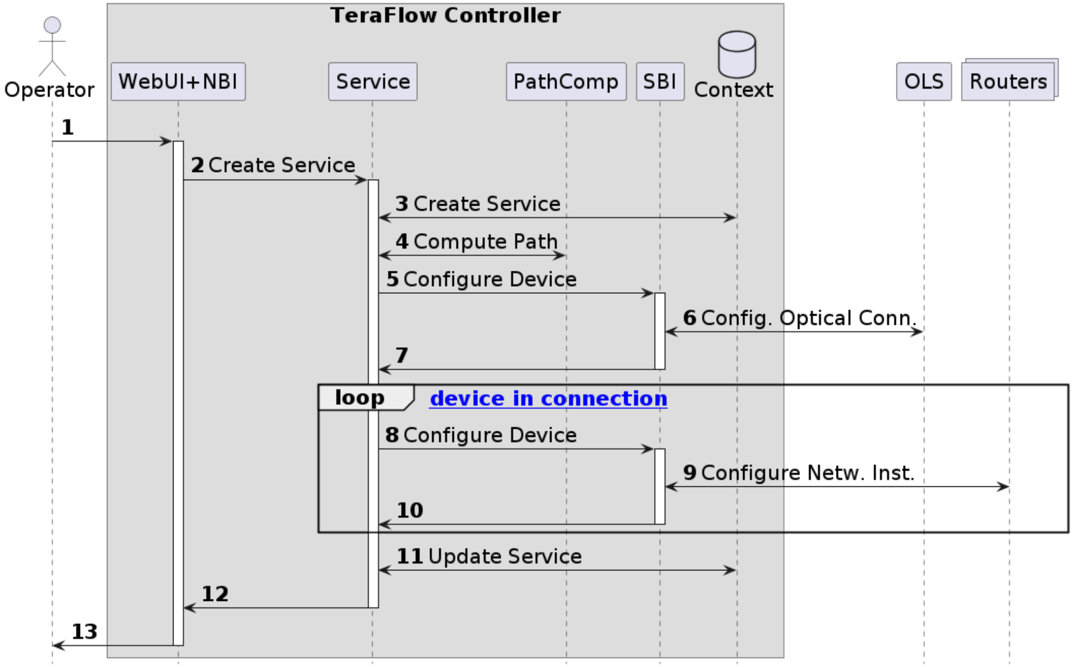
Programmable and flexible network configuration

The core TeraFlowSDN microservices that are relevant to this work:

- **Context component** is a datastore for key informational entities (i.e., network topologies, devices, etc.).
- **SBI** manages the underlying network equipment both for configuring them and for collecting monitoring data from them.
- **Service component** oversees the lifecycle of connectivity services within the network.
- **PathComp component** performs path computations and infers the appropriate configuration rules they require.
- **Monitoring component** handles the collection of telemetry data from the network elements through the SBI and produces usable monitoring samples.
- **NBI** serves as the communication bridge between TeraFlowSDN and external OSS/BSS system.
- **Web UI component** provides a user-friendly interface for the network operators.



TeraFlowSDN Micro-service-based Architecture



Create Service Workflow in TeraFlowSDN

Role in the PoC

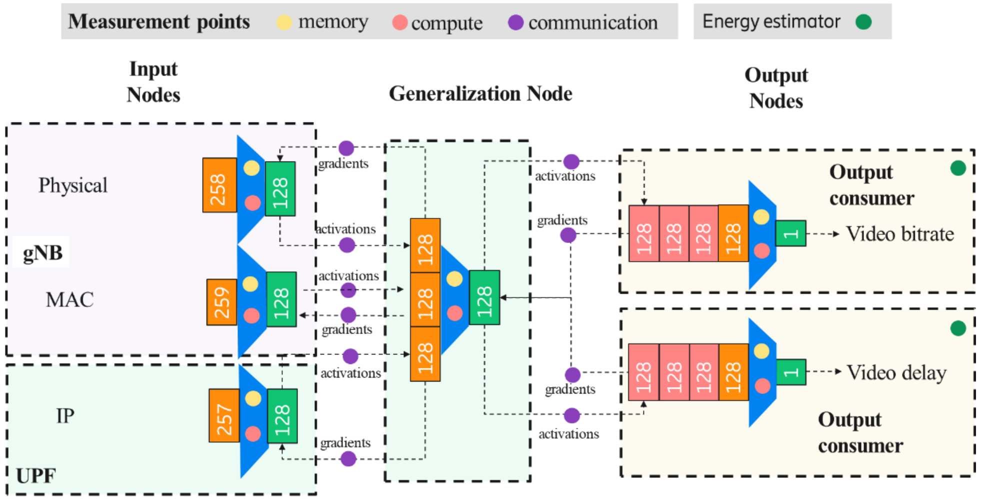
- Provides connectivity services upon a backhaul Transport Network

Training and inference of collaborative distributed machine learning model on a dynamically changing heterogeneous 6G architecture environment

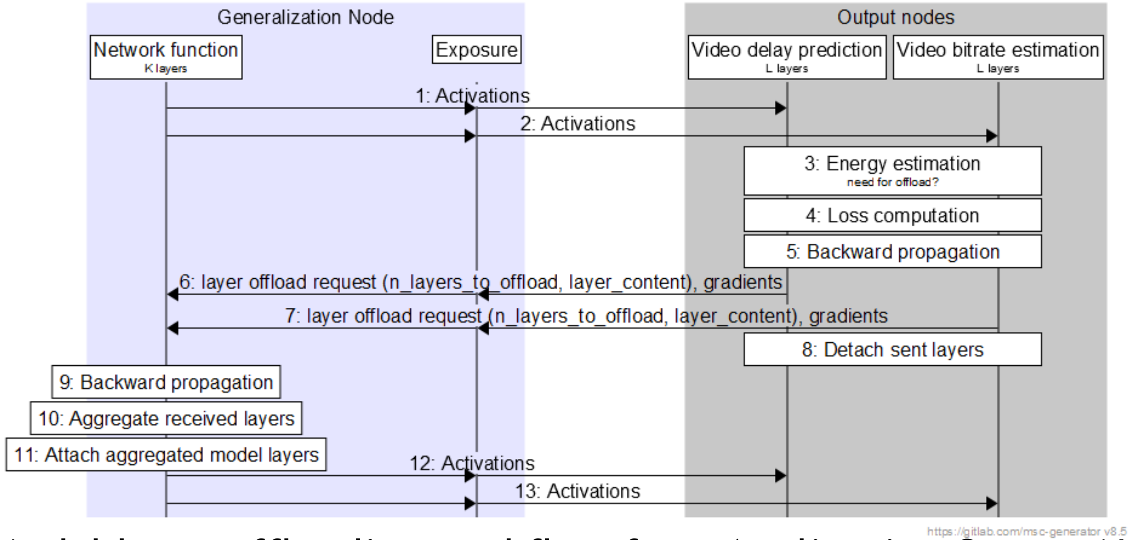
Collaborative training across application and core network allowing a joint Neural Network (NN) model to be trained for a video streaming quality estimation task. Based on split training that enables:

- Cross-domain training and inference, without necessitating to move dataset between distributed data nodes
- Model generalization that learns common representation of data to serve multi-tasks (use cases)
- Model layer offloading when computation and energy resources are scarce.

It facilitates offloading NN model layers from a distributed node to adapt to the available compute, memory and energy.



Distributed model architecture



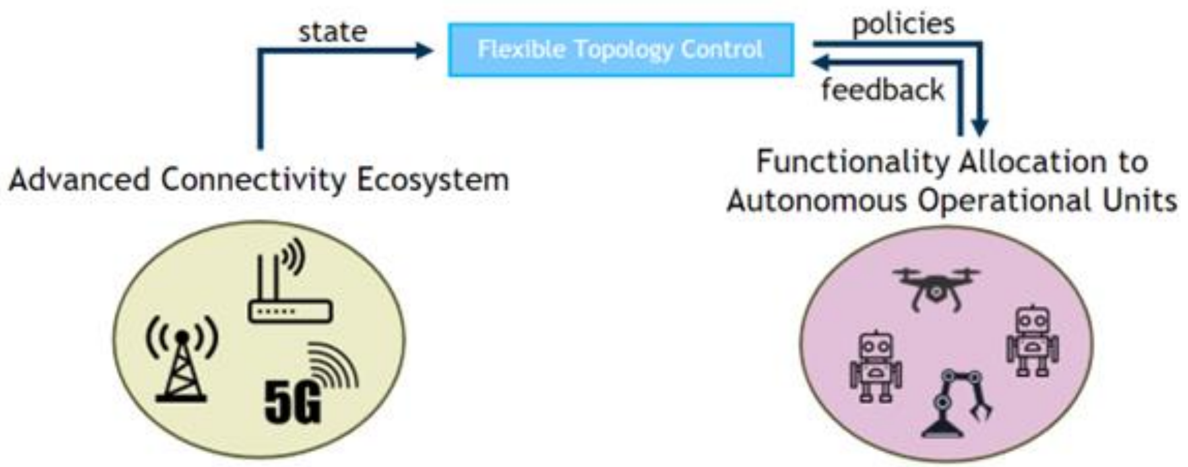
Model layer offloading workflow from Application Output Node to the Network Function in the Core Network

Role in the PoC:

- It is expected to have impact on the following enablers: AI-related functionalities in the 6G system blueprint such as video streaming, data exposure, and AIaaS.

Trustworthy flexible topologies and beyond communication aspects

- Flexible topologies are designed to adapt quickly to resource and coverage constraints, making them essential for addressing the dynamic needs of 6G environments.
- As described in D3.3, these networks utilize components such as node discovery, trust/cost evaluation, and AI/ML-driven resource optimization to form adaptable and efficient network structures.
- A Flexible Topology Node (FTN) is implemented on an autonomous UAV with multi-connectivity capabilities:
  - The FTN communicates with an edge server via a 5G network interface to maintain up-to-date information on Worker Nodes' (WNs) locations and tasks.
  - Within System-PoC B activities, two scenarios trigger dynamic network allocation using the FTN: local connectivity loss and out-of-coverage intent.


state
Flexible Topology Control
policies
feedback
Advanced Connectivity Ecosystem
Functionality Allocation to Autonomous Operational Units
5G

Role in the PoC:

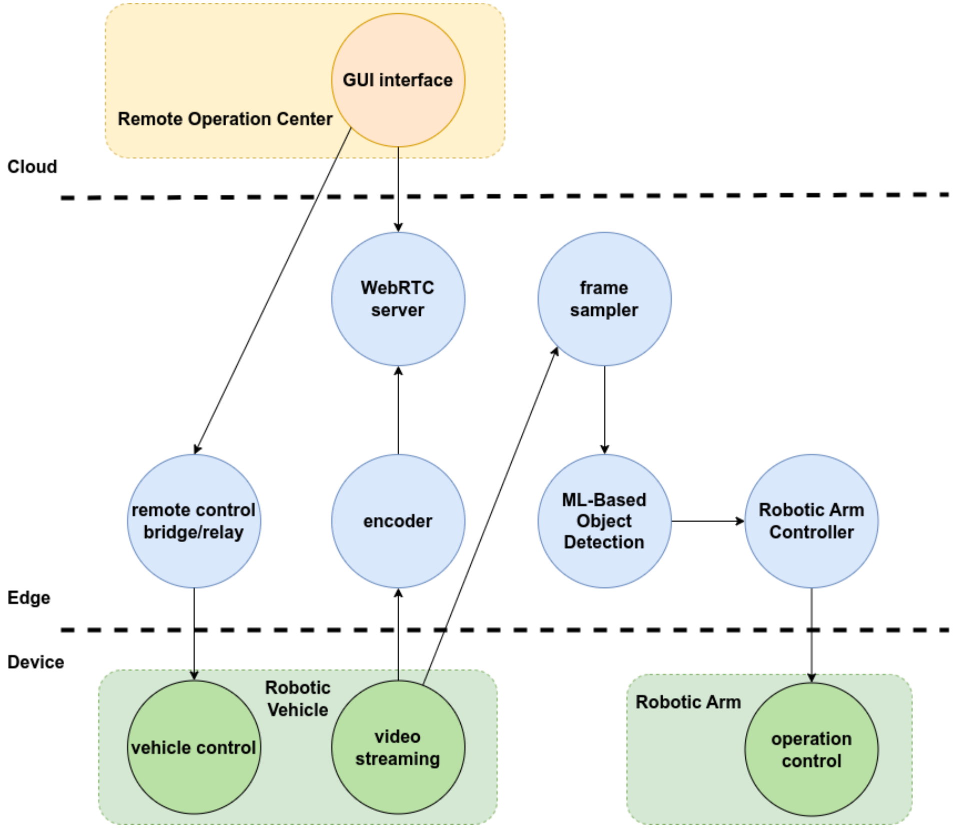
- To ensure state and policy feedback loops that enhance network performance and operational efficiency.

AI-assisted E2E lifecycle management of a 6G latency-sensitive service across the compute continuum

Component that focuses on real-time monitoring of manufacturing processes and ensuring safety through prompt responses to detected anomalies.

Offered mechanisms and functionalities:

- **Service autoscaling and migration:** To horizontally scale and migrate the application from the edge to the cloud (& vice-versa) based on the measured workload and E2E latency.
- **DLT-based service federation:** To migrate latency-sensitive applications from the consumer to the provider domain using a blockchain smart contract to agree on the federation terms.
- **Extreme-edge cluster emulation on high availability scenarios:** To demonstrate high availability scenarios on top of an emulated extreme-edge infrastructure.
- **Inter-cluster communications scenarios:** Two alternative options are explored.
  - karmada multi-cluster management tool
  - Network Service Mesh (NSM)



6G-sensitive service application graph description

E2E Implementation Scenarios

Autonomous operation in inventory management

Implementation scenario of the *Automated Inventory Management* solution addresses the following workflow:

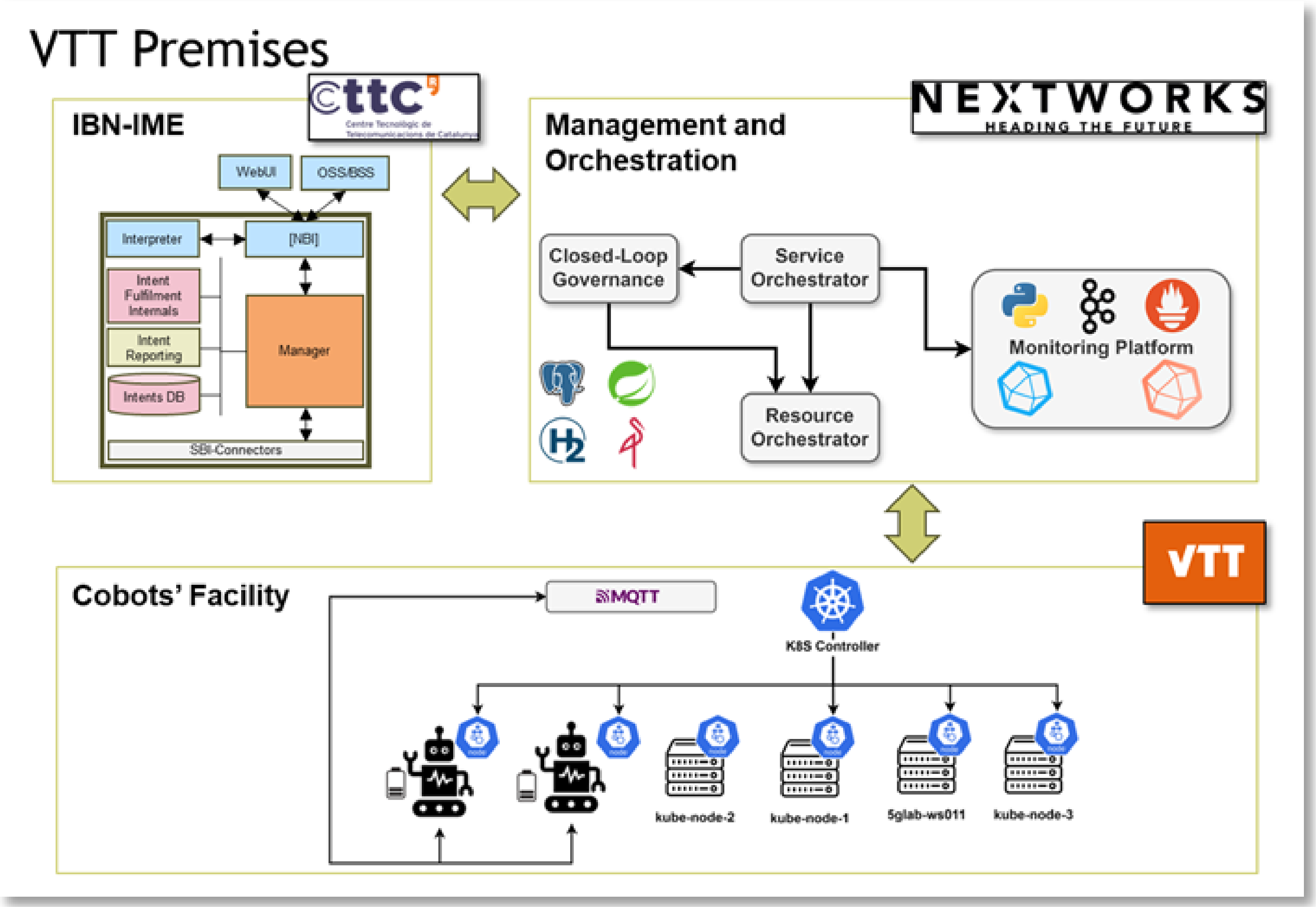
1) An infrastructure monitoring component to continuously monitor various KPIs and the status of physical and virtual resources within the system.
   - If any metric exceeds a predefined threshold or a device becomes unavailable, the infrastructure monitoring component triggers an alert to the API server.
2) The API server receives intent-based requests, i.e., high-level commands, from end-users or external sources.
   - These requests express specific needs or objectives to be accomplished by the system.
3) Upon receiving a request from either the infrastructure monitoring component or an external intent-based request, the API server triggers the functionality allocation mechanism.
   - It is responsible for determining the optimal allocation of resources to meet the request's requirements.
4) The API server acts as an intermediary, gathering the requested information from the infrastructure monitoring component for resource status and capabilities; the service registry for task requirements and workload information; and the trust evaluation function for trust indices.
5) The functionality allocation mechanism then performs an optimisation process using the gathered data.
   - Based on the optimisation process, a proposed placement solution is generated and sent to the API server.
6) The API server verifies the feasibility of the solution and then forwards the decided action to the orchestration enforcer.
   - A placement decision is implemented on the system's resources.
7) Once the orchestration enforcer completes the placement, it sends a confirmation response back to the API server.
8) The API server then updates the infrastructure monitoring component and the service registry with the new placement information, ensuring the system's state is current and reflective of the latest changes.

Zero-touch Cobot-based video surveillance

Objective:
Request the deployment of a cobot (i.e., collaborative robot) surveillance service using intent-based requests that trigger the deployment and an associated closed-loop to manage and monitor the deployed service.

Service Provisioning

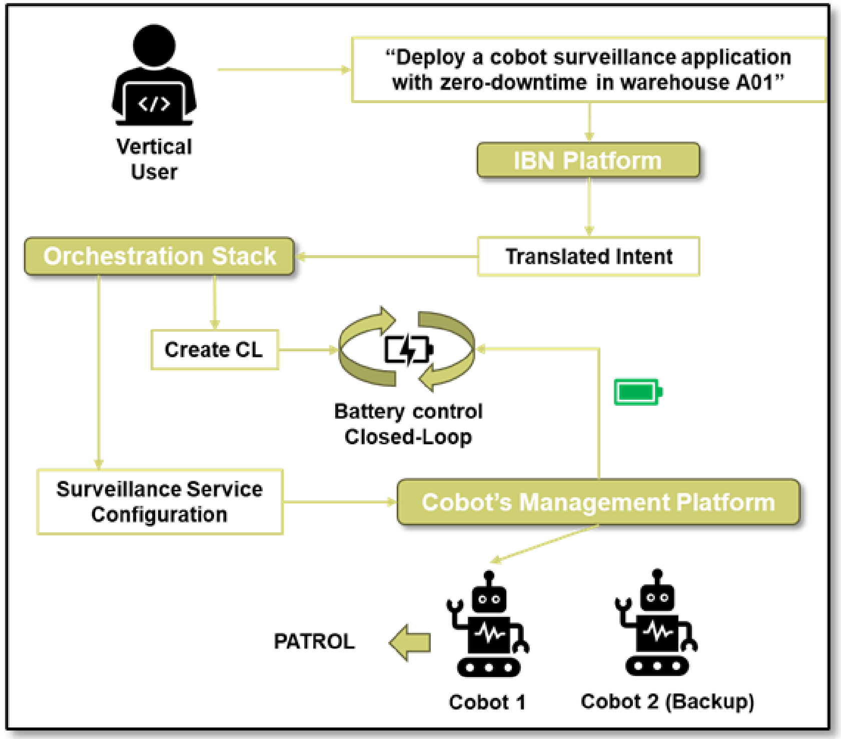
- IBN-IME (CTTC)
  - Intent-based Service LCM
- Management and Orchestration (NXW)
  - Orchestration
  - Monitoring
  - CL Governance
- Cobots' Facility (VTT)
  - Cobots platform
  - Edge/Extreme-Edge continuum

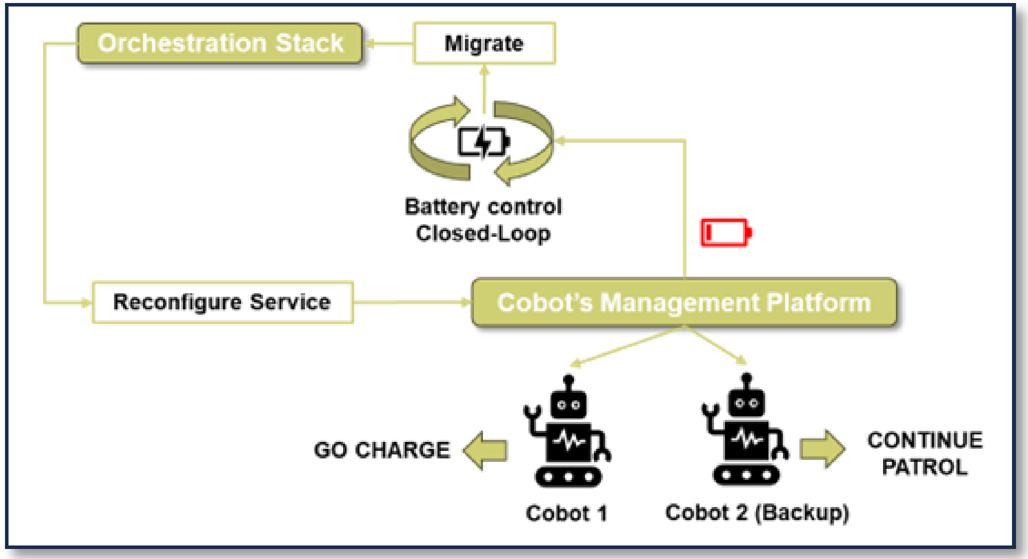


Zero-touch Cobot-based video surveillance workflow

Service Provisioning

- A Vertical User request a cobot-based surveillance service by specifying an intent on the IBN Platform
- The IBN Platform parses and translate the intent in requests for the Orchestration Stack
- The Orchestration Stack provision the service on the cobot's Management Platform
- In parallel specific Closed-Loop (CL) are deployed to guarantee the service continuity in case of cobot's battery drain


Vertical User
"Deploy a cobot surveillance application with zero-downtime in warehouse A01"
IBN Platform
Translated Intent
Orchestration Stack
Create CL
Battery control Closed-Loop
Surveillance Service Configuration
Cobot's Management Platform
PATROL
Cobot 1
Cobot 2 (Backup)



Closed-Loop Automation

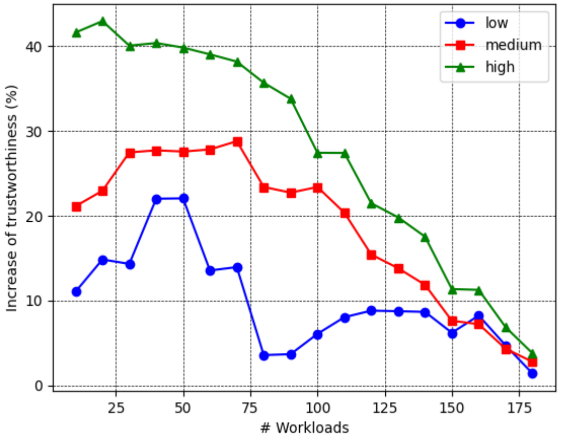
- When battery charge drops below 15%, the CL enforce the migration of the service towards another cobot

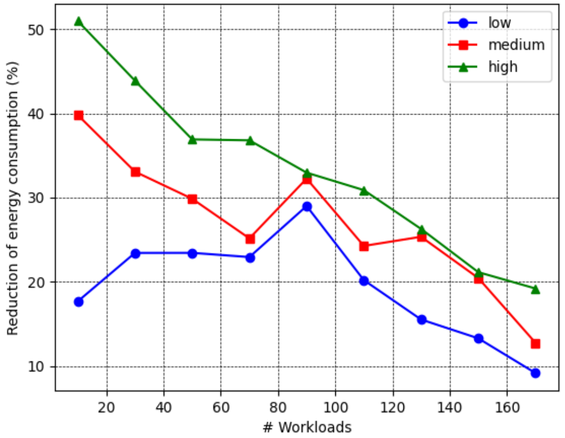
System-PoC #B evaluation results for sustainable and trustworthy 6G systems

Autonomous Operation Scenario:
Advanced M&O, Flexible topologies and Network beyond communications enablers in Cobot-powered Warehouse Inventory Management (1)

- Preliminary results obtained related to trustworthiness (left graph) and energy consumption (right graph) with the utilisation of the functionality allocation algorithm based on the genetic algorithm paradigm, compared to the feasible round-robin placement algorithm (baseline).

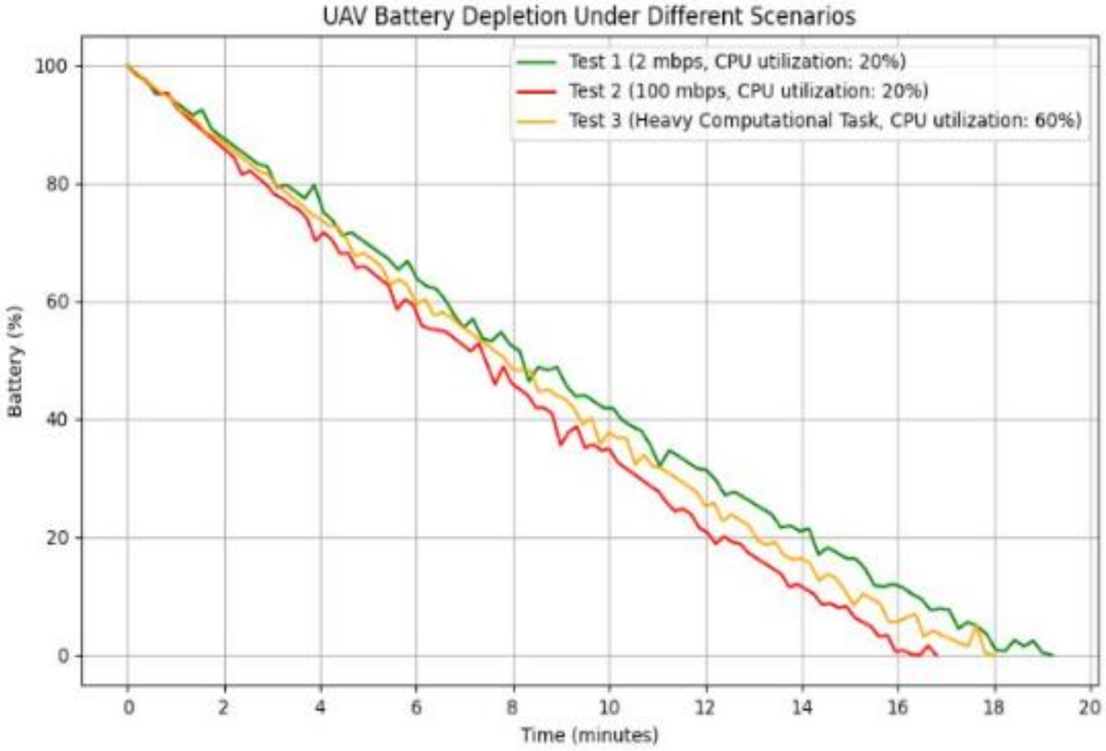
- The measurements taken with different trust weight and energy weight levels compared to the other terms of the function, respectively:
  - low (almost zero contribution)
  - medium (moderate contribution)
  - high (full contribution)





UAV battery test scenarios characteristics

Conf. #	Setup
1	Transmission Load: 2Mbps, CPU Utilization: 22%, CPU cores: 7, Battery Autonomy: 0.32h
2	Transmission Load: 100Mbps, CPU Utilization: 22%, CPU cores: 7, Battery Autonomy: 028h
3	Transmission Load: 0Mbps, CPU Utilization: 60%, CPU cores: 7, Battery Autonomy: 0.30h

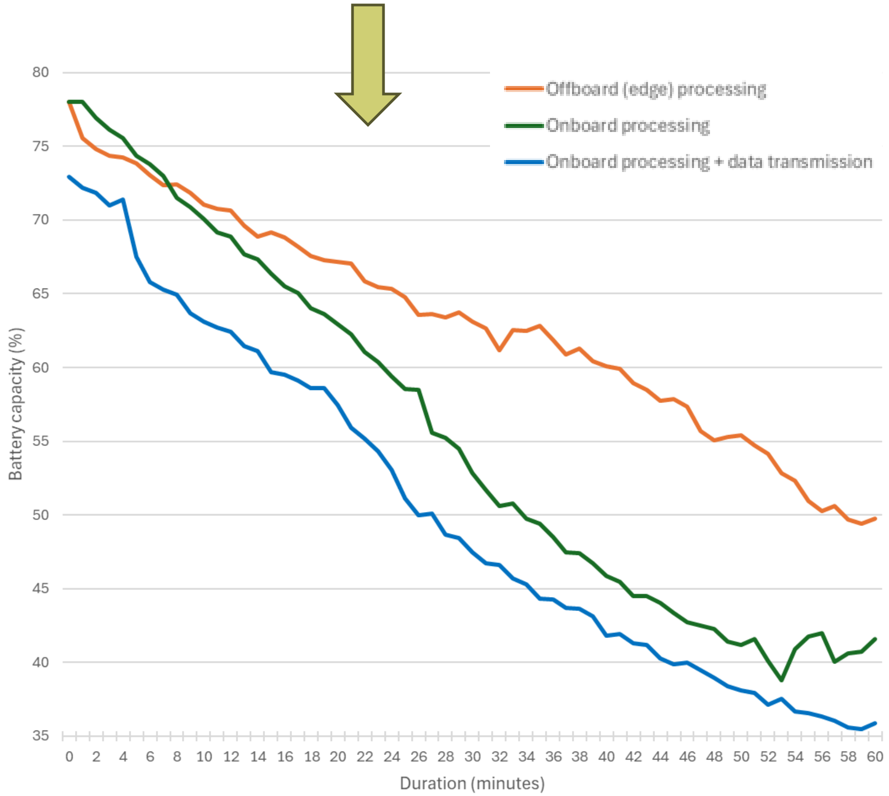


Autonomous Operation Scenario: Advanced M&O, Flexible topologies and Network beyond communications enablers in Cobot-powered Warehouse Inventory Management (2)

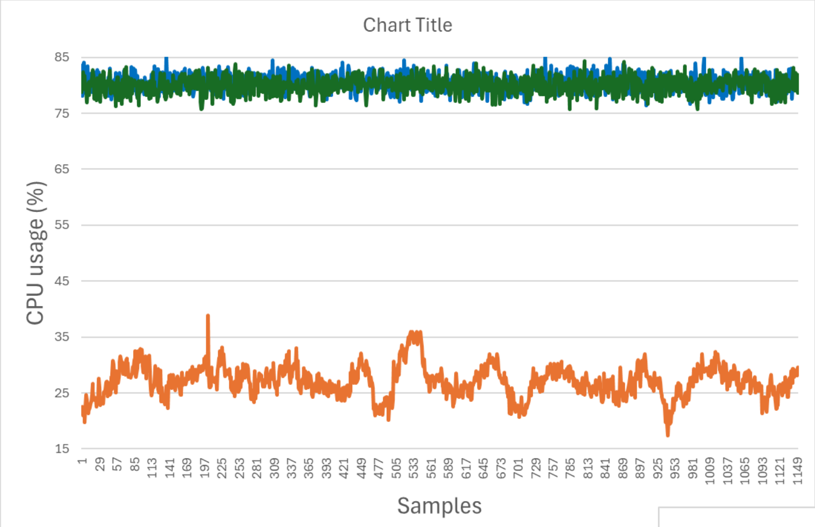
Power consumption aspects when the AI/ML operations are allocated:

- at the edge (compute offloading)
- at the robotic platform (on-board processing).

→ Reduced battery depletion of approximately 20% under heavy ML operations.

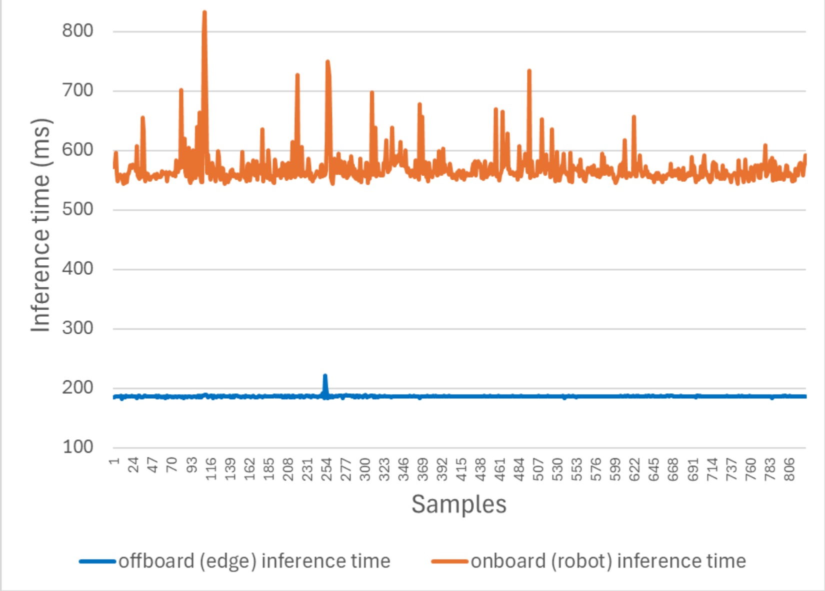


The effect of offloading ML operations was also examined, with regards to CPU utilization and inference time of the ML model.



Performance benefits of offloading computation via 5G on CPU utilisation

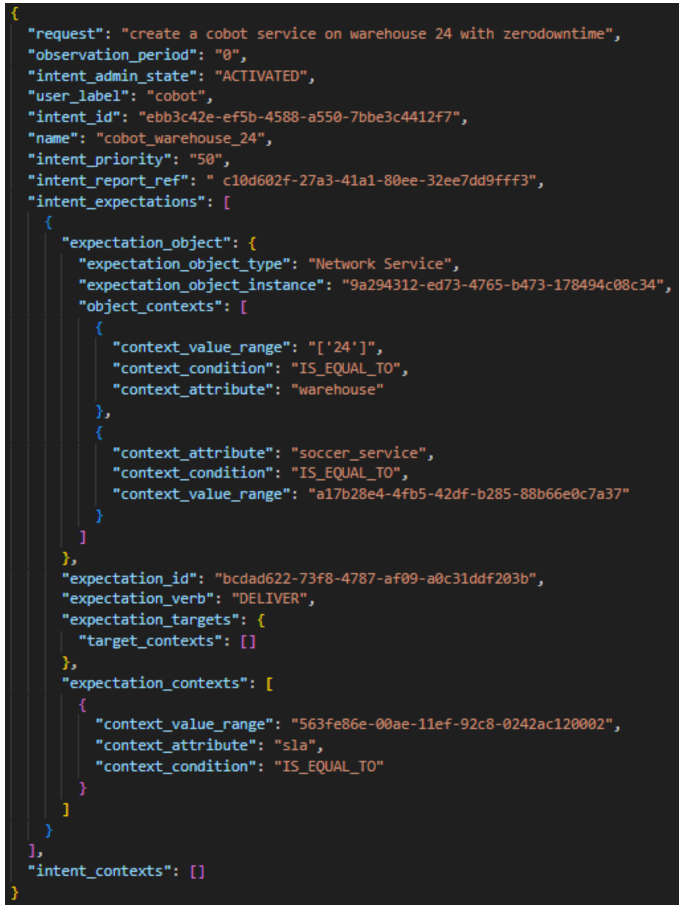
Performance benefits of offloading computation via 5G on inference time analysis



Zero-touch Cobot-based video surveillance Scenario: Intent-based Service Provisioning Results (1)

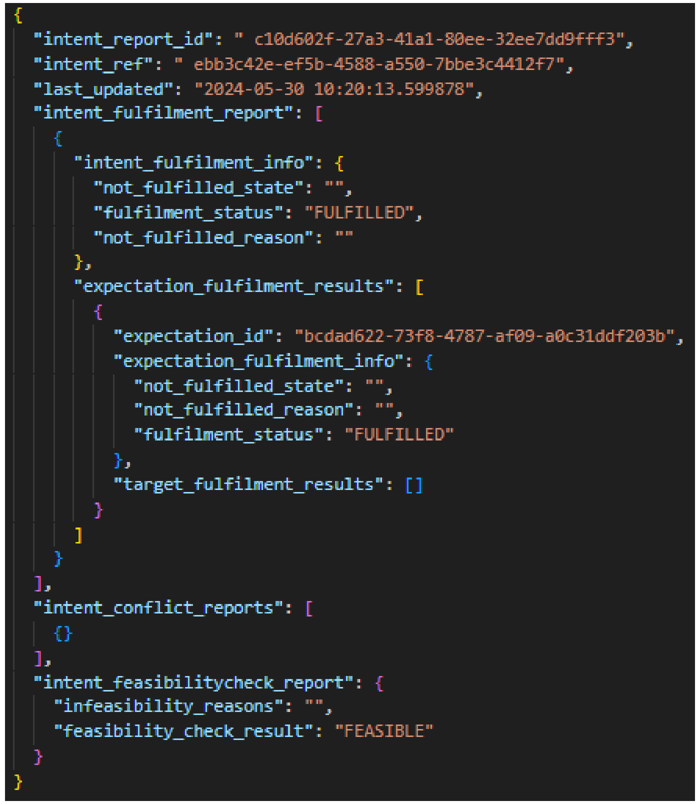
"create a cobot service on warehouse 24 with zerodowntime"

Intent data object



```json
{
  "request": "create a cobot service on warehouse 24 with zerodowntime",
  "observation_period": "0",
  "intent_admin_state": "ACTIVATED",
  "user_label": "cobot",
  "intent_id": "ebb3c42e-ef5b-4588-a550-7bbe3c4412f7",
  "name": "cobot_warehouse_24",
  "intent_priority": "50",
  "intent_report_ref": " c10d602f-27a3-41a1-80ee-32ee7dd9fff3",
  "intent_expectations": [
    {
      "expectation_object": {
        "expectation_object_type": "Network Service",
        "expectation_object_instance": "9a294312-ed73-4765-b473-178494c08c34",
        "object_contexts": [
          {
            "context_value_range": "['24']",
            "context_condition": "IS_EQUAL_TO",
            "context_attribute": "warehouse"
          },
          {
            "context_attribute": "soccer_service",
            "context_condition": "IS_EQUAL_TO",
            "context_value_range": "a17b28e4-4fb5-42df-b285-88b66e0c7a37"
          }
        ]
      },
      "expectation_id": "bcdad622-73f8-4787-af09-a0c31ddf203b",
      "expectation_verb": "DELIVER",
      "expectation_targets": {
        "target_contexts": []
      },
      "expectation_contexts": [
        {
          "context_value_range": "563fe86e-00ae-11ef-92c8-0242ac120002",
          "context_attribute": "sla",
          "context_condition": "IS_EQUAL_TO"
        }
      ]
    }
  ],
  "intent_contexts": []
}
```

Intent Report data object

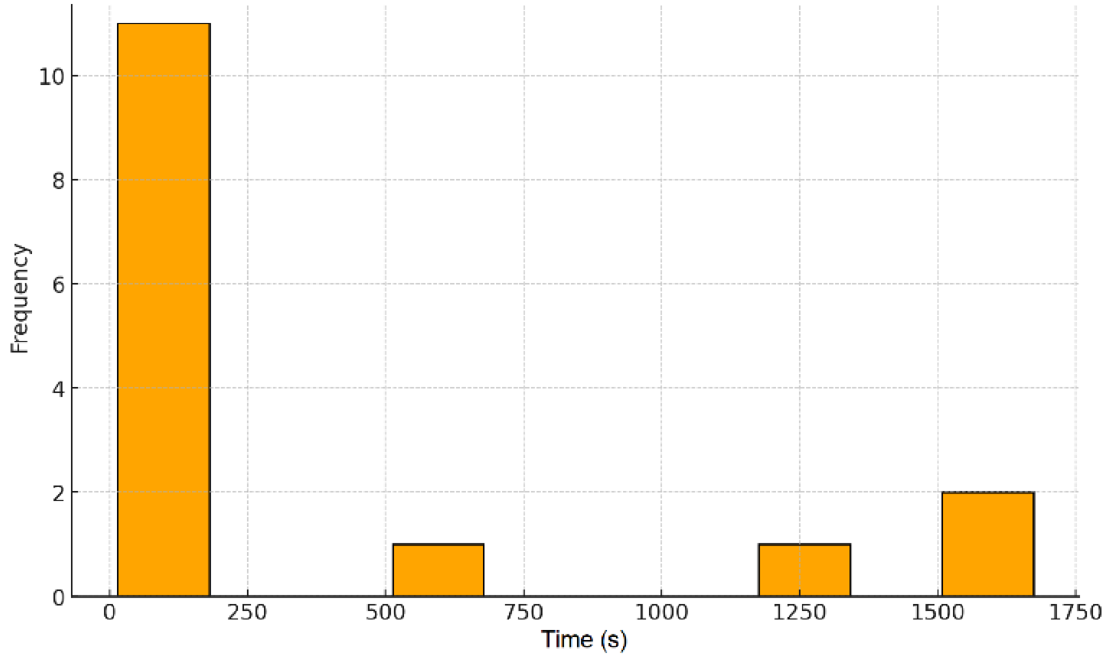


```json
{
  "intent_report_id": " c10d602f-27a3-41a1-80ee-32ee7dd9fff3",
  "intent_ref": " ebb3c42e-ef5b-4588-a550-7bbe3c4412f7",
  "last_updated": "2024-05-30 10:20:13.599878",
  "intent_fulfilment_report": [
    {
      "intent_fulfilment_info": {
        "not_fulfilled_state": "",
        "fulfilment_status": "FULFILLED",
        "not_fulfilled_reason": ""
      },
      "expectation_fulfilment_results": [
        {
          "expectation_id": "bcdad622-73f8-4787-af09-a0c31ddf203b",
          "expectation_fulfilment_info": {
            "not_fulfilled_state": "",
            "not_fulfilled_reason": "",
            "fulfilment_status": "FULFILLED"
          },
          "target_fulfilment_results": []
        }
      ]
    }
  ],
  "intent_conflict_reports": [
    {}
  ],
  "intent_feasibilitycheck_report": {
    "infeasibility_reasons": "",
    "feasibility_check_result": "FEASIBLE"
  }
}
```

Set of tests: 15

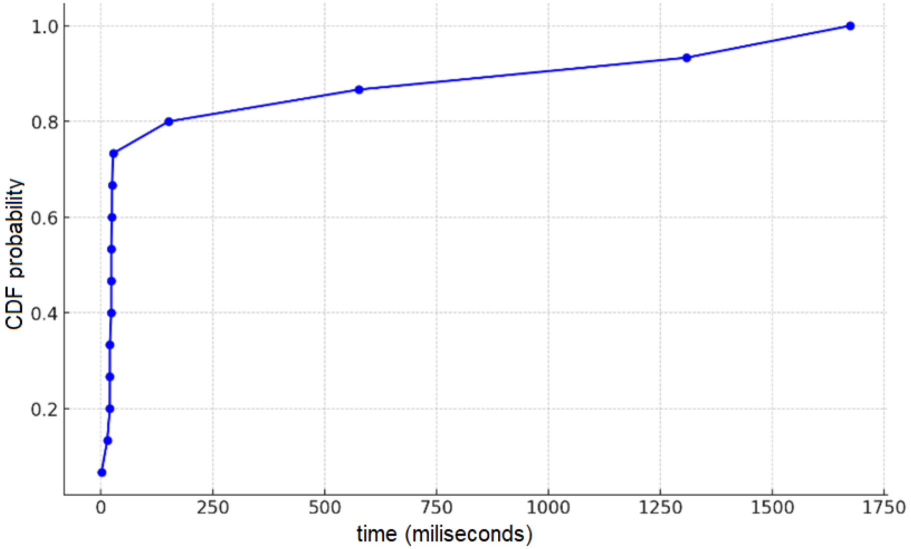
11 of tests were deployed with less than 250s

4 of the test required longer time.

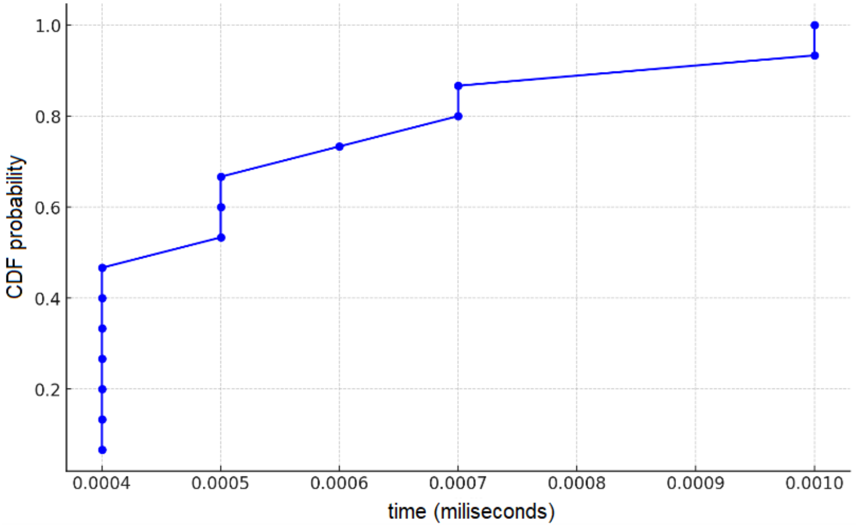


Zero-touch Cobot-based video surveillance Scenario: Intent-based Service Provisioning Results (2)

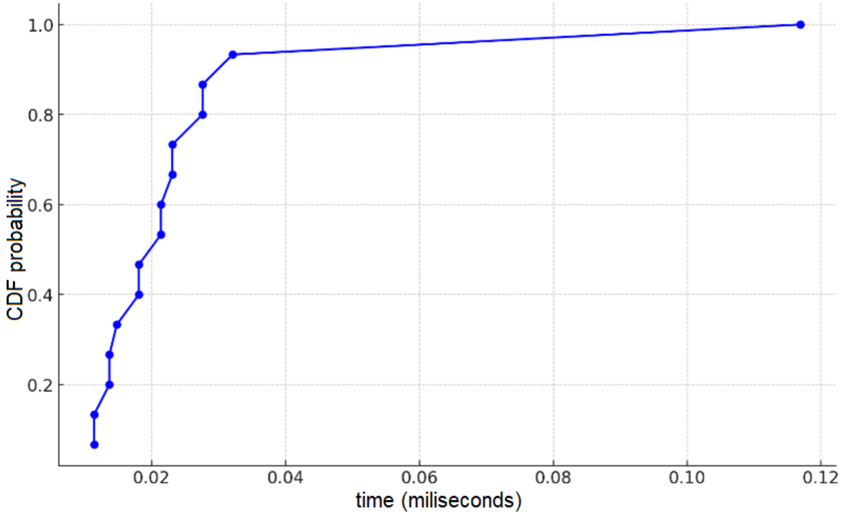
Vertical service instance creation process CDF.



Intent feasibility process CDF.



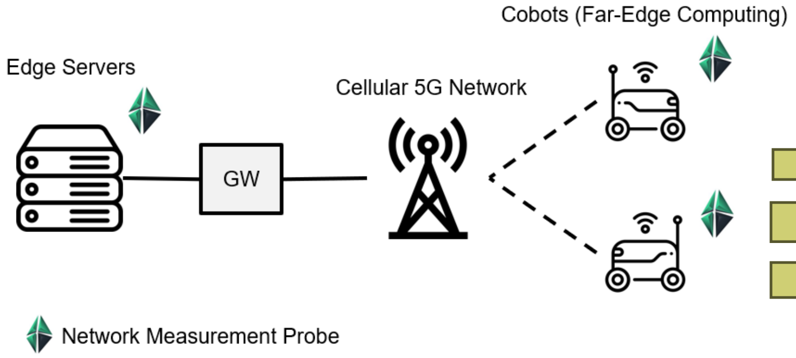
Natural Language Processing (NLP) process CDF.



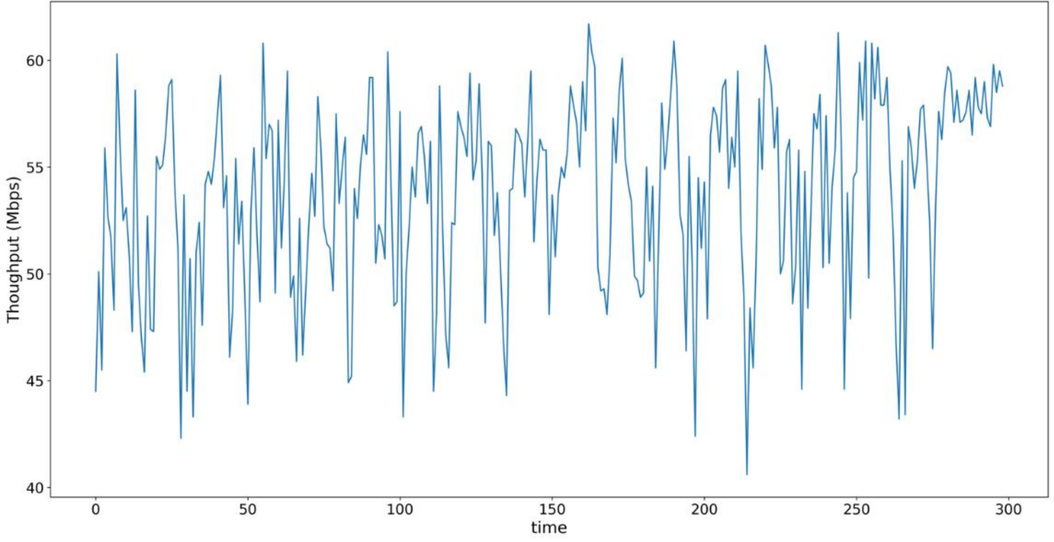
Initial Conclusions:

- Creating the vertical instance takes most of the time.
- Applying NLP does not add a big amount of time to the overall service provisioning.

Zero-touch Cobot-based video surveillance Scenario: Integrated Closed-Loop for Cobot Service Migration



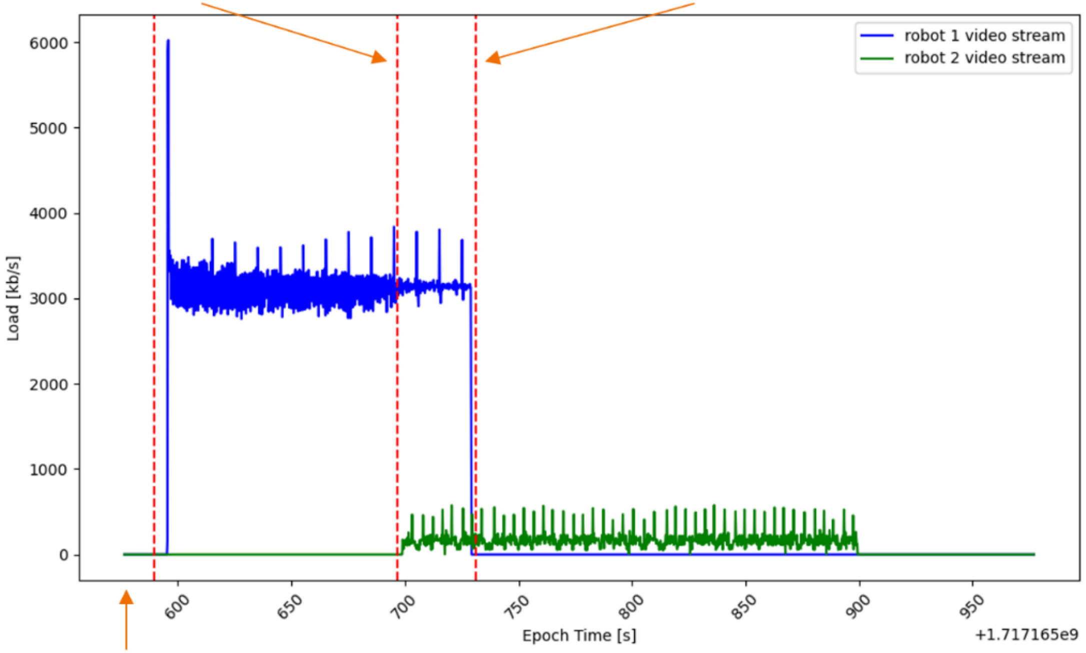
Cobots Uplink throughput in a 5G SA network.



Performance evaluation of a surveillance service migration.

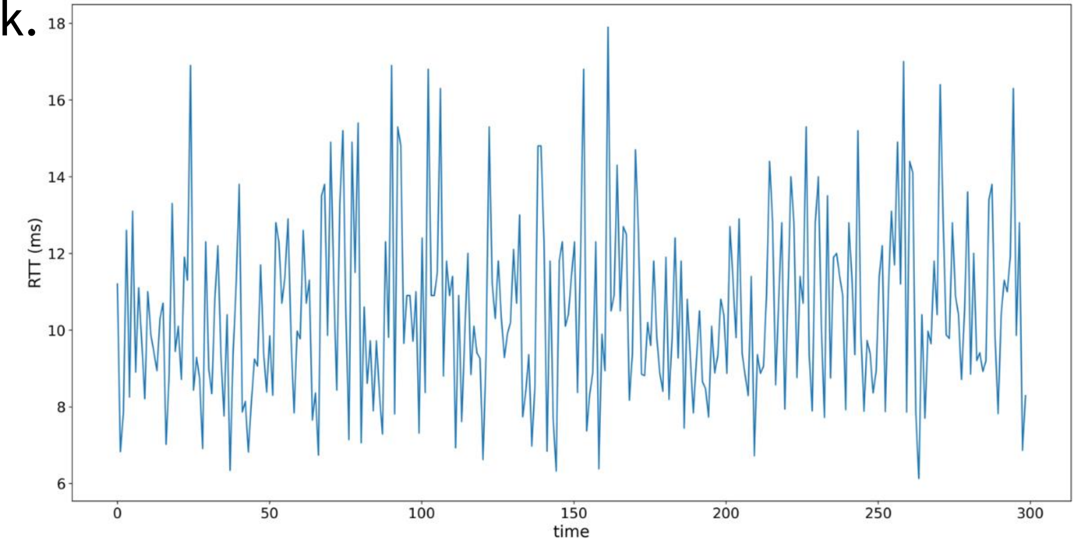
Replacement is triggered

Video from robot-2 is visible



Start the demo

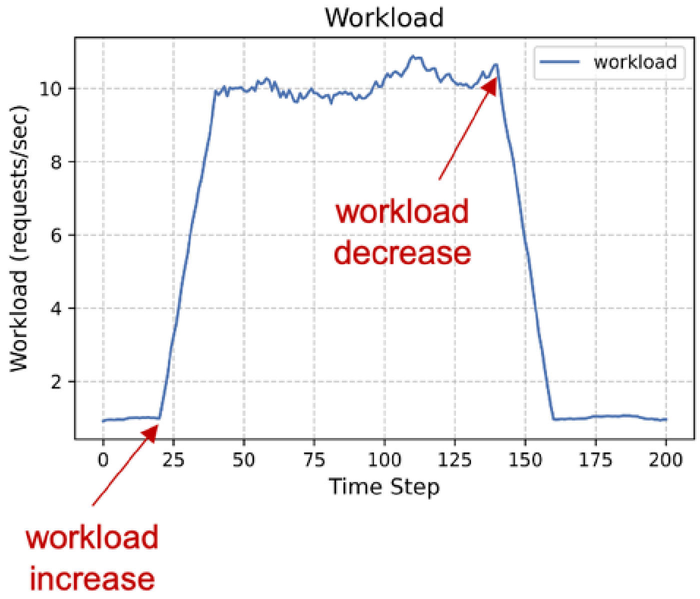
Cobots Round-Trip (ms) to the Edge Servers in a 5G SA network.

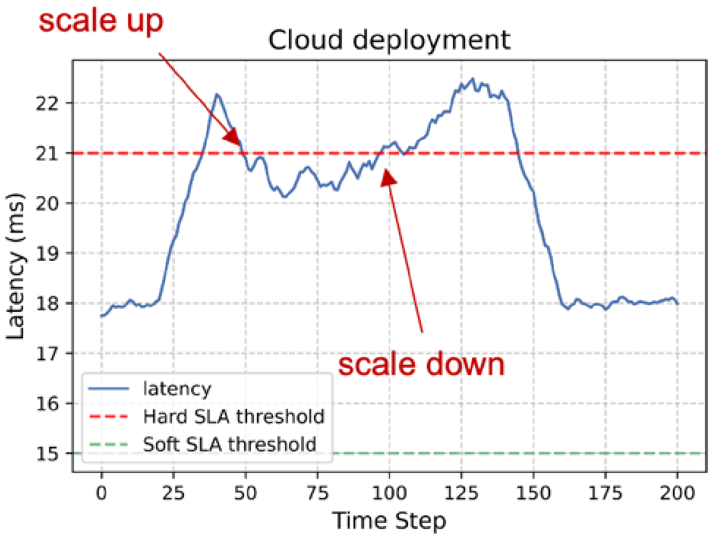


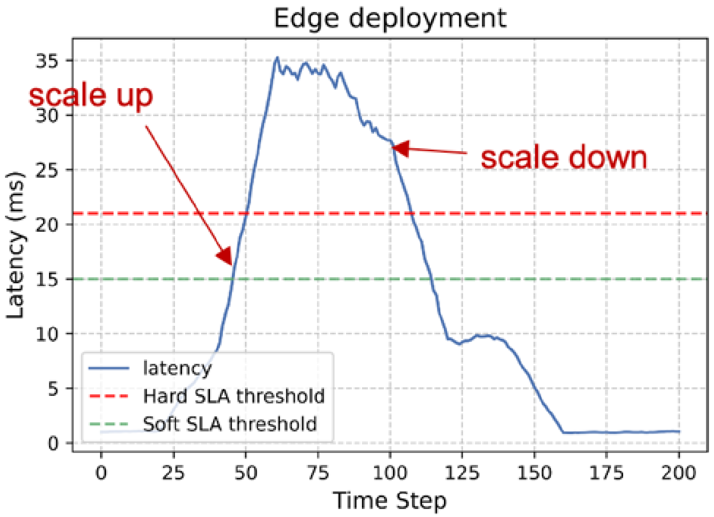
PoC Component:
AI-assisted E2E lifecycle management of a 6G latency-sensitive service across the compute continuum (1)

- Two different clusters are used to experiment how horizontal scaling influences performance:
  - Cloud deployment: When the service is executed at the cloud, scaling up successfully reduces latency by 10%. → Requests sent to the cloud suffer from additional communication delays (~15ms).
  - Edge deployment: At the Edge, communication latency is lower, but scaling up actually increases latency (up to 3100%) → Not enough CPU available and concurrency hinders performance severely.

- A hybrid deployment can tackle such a trade-off, working at the Edge during normal operation times, and offloading high workloads to the cloud where scaling up is efficient.



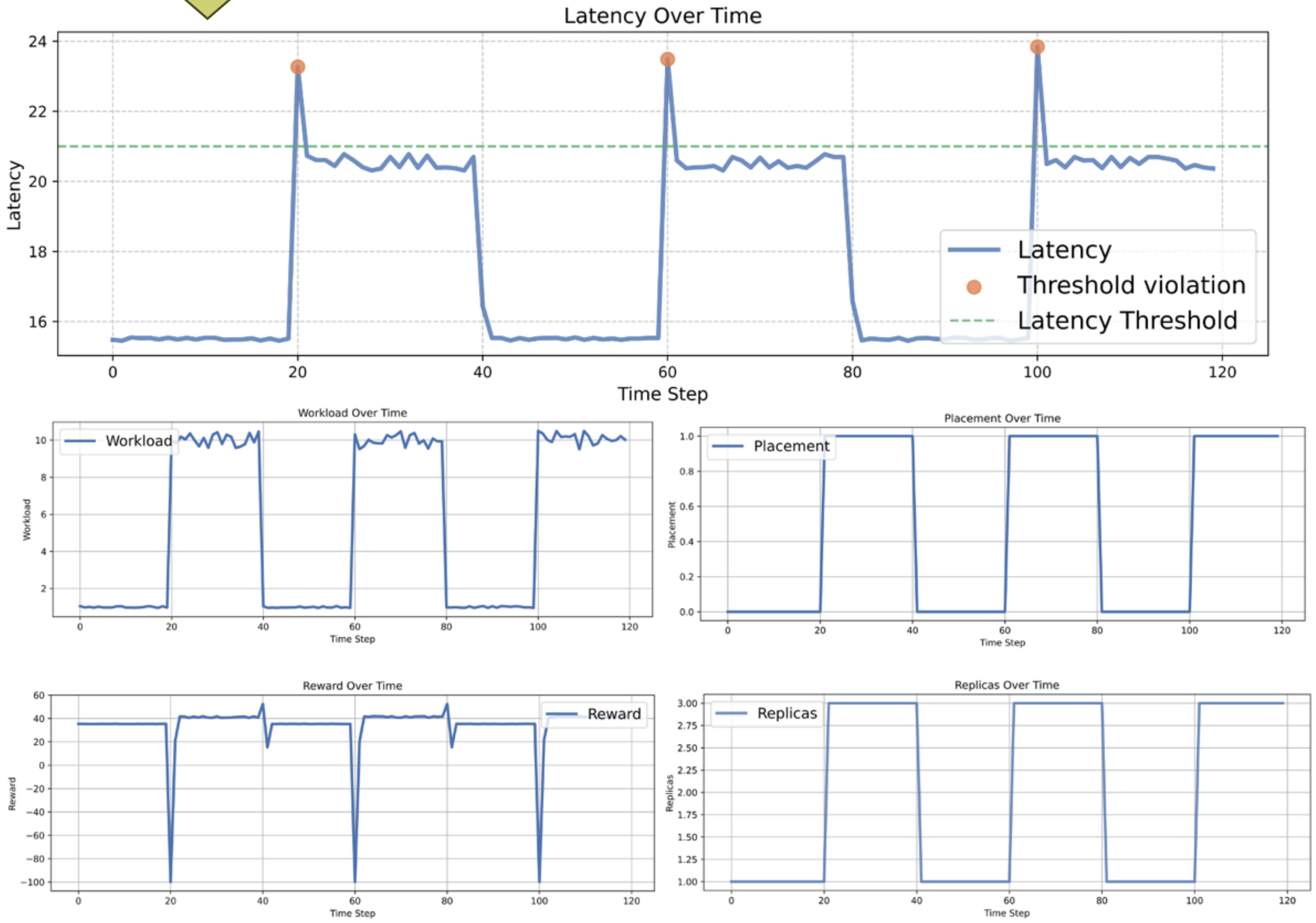


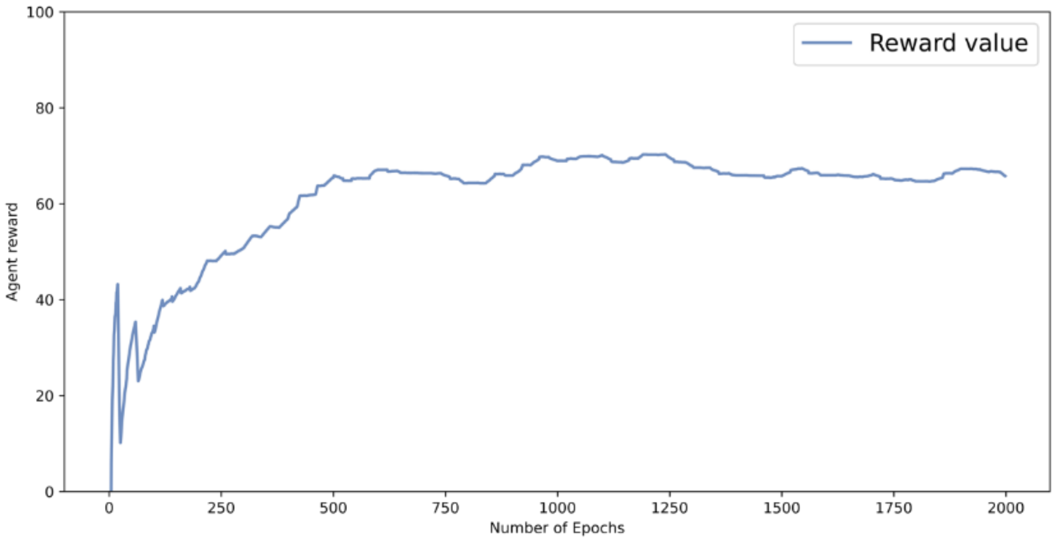


PoC Component:
AI-assisted E2E lifecycle management of a 6G latency-sensitive service across the compute continuum (1)

- RL agent autoscaling performance: workload, decisions, SLO satisfaction and agent reward.

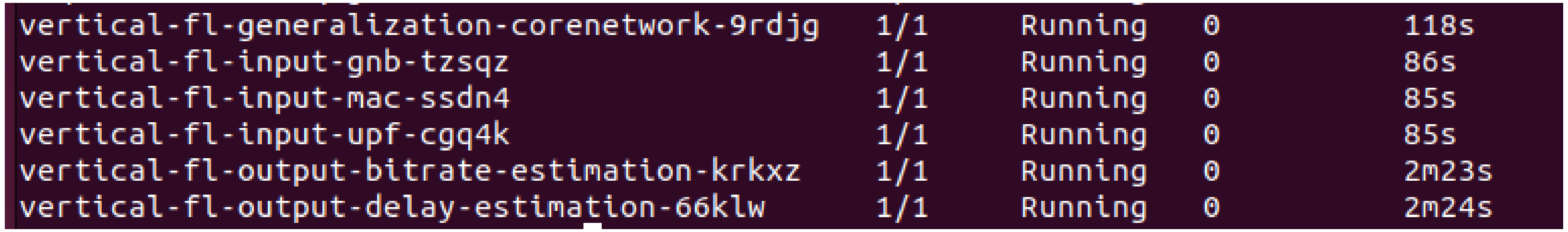
- RL autoscaling agent training: reward evolution across the first 200 episodes





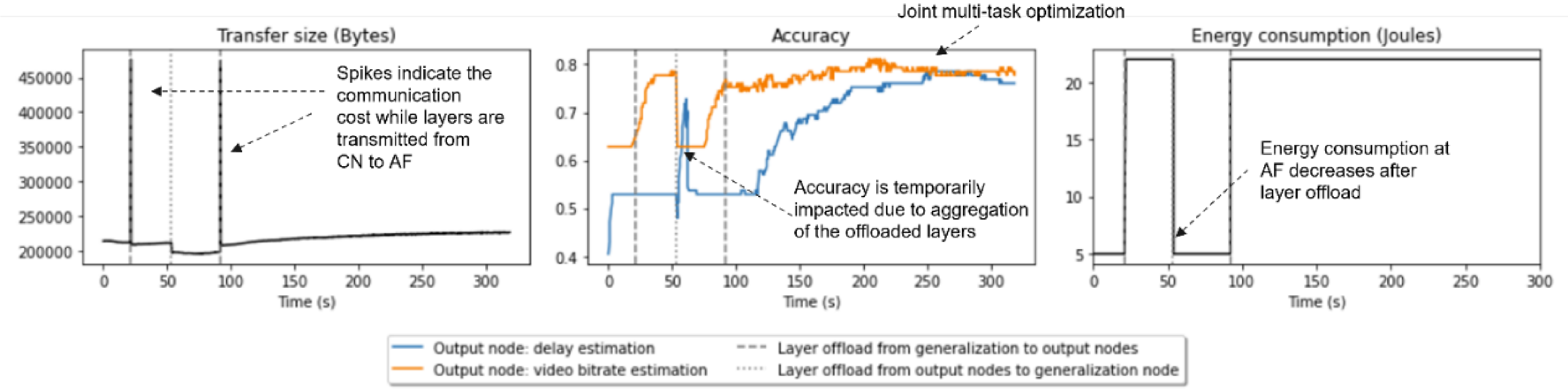
PoC Component:
Training and inference of collaborative distributed machine learning model on a dynamically changing heterogeneous 6G architecture environment

- This System-PoC #B Component consists of 3 Kubernetes pods for input nodes, 1 pod for generalization node, and 2 pods for output nodes.



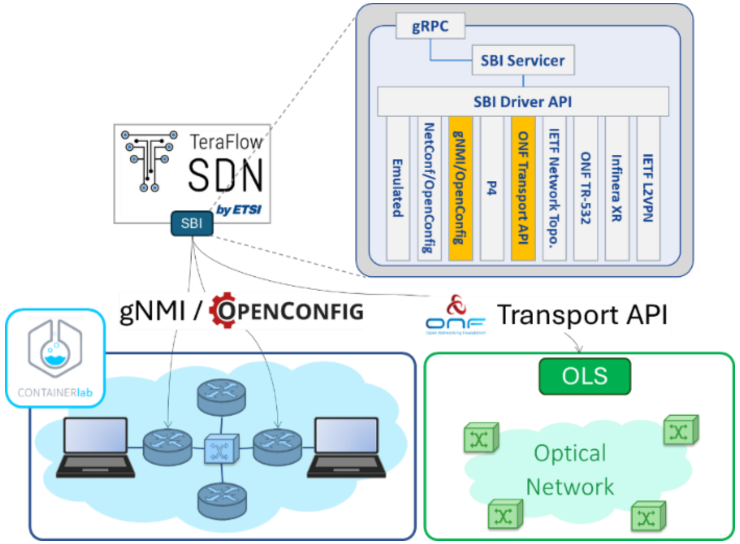
```
vertical-fl-generalization-corenetwork-9rdjg   1/1     Running   0          118s
vertical-fl-input-gnb-tzsqz                    1/1     Running   0          86s
vertical-fl-input-mac-ssdn4                    1/1     Running   0          85s
vertical-fl-input-upf-cgq4k                    1/1     Running   0          85s
vertical-fl-output-bitrate-estimation-krkxz    1/1     Running   0          2m23s
vertical-fl-output-delay-estimation-66klw      1/1     Running   0          2m24s
```

- Transferred information size from generic node to the output nodes (left), accuracy observed at the output nodes (middle), and the estimated energy consumption at the output nodes (right) are given.

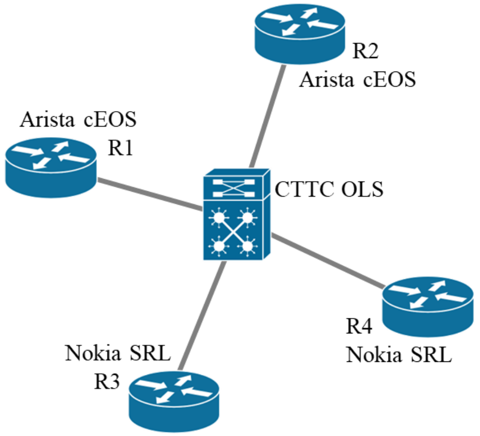


PoC Component:
Programmable and flexible network configuration

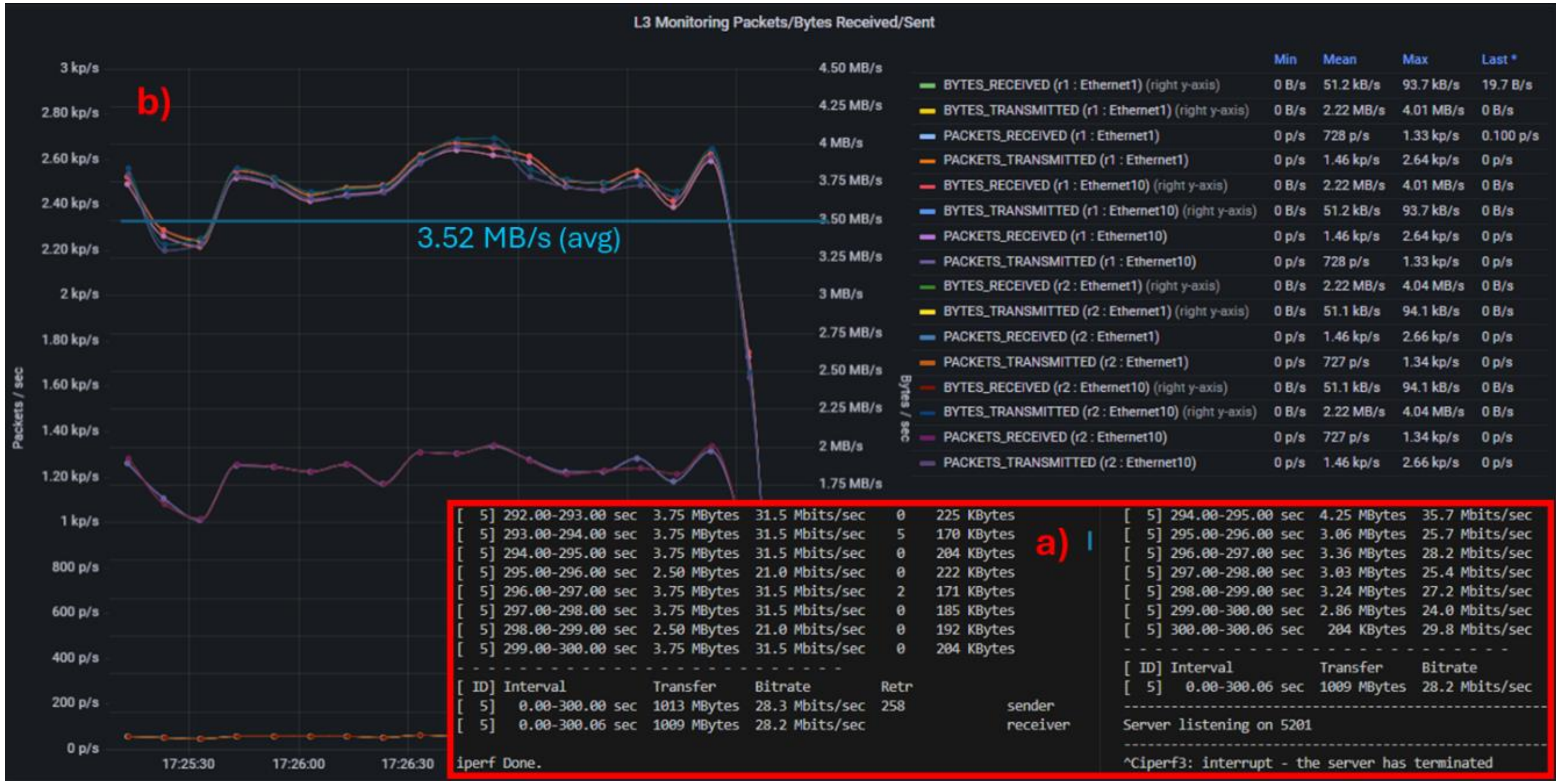
- Testbed Architecture for DataPlane-in-a-Box.



- Network Topology in TeraFlowSDN



- Connectivity (a) and Performance Measurement (b) using DataPlane-in-a-box


L3 Monitoring Packets/Bytes Received/Sent

	Min	Mean	Max	Last *
BYTES_RECEIVED (r1 : Ethernet1) (right y-axis)	0 B/s	51.2 kB/s	93.7 kB/s	19.7 B/s
BYTES_TRANSMITTED (r1 : Ethernet1) (right y-axis)	0 B/s	2.22 MB/s	4.01 MB/s	0 B/s
PACKETS_RECEIVED (r1 : Ethernet1)	0 p/s	728 p/s	1.33 kp/s	0.100 p/s
PACKETS_TRANSMITTED (r1 : Ethernet1)	0 p/s	1.46 kp/s	2.64 kp/s	0 p/s
BYTES_RECEIVED (r1 : Ethernet10) (right y-axis)	0 B/s	2.22 MB/s	4.01 MB/s	0 B/s
BYTES_TRANSMITTED (r1 : Ethernet10) (right y-axis)	0 B/s	51.2 kB/s	93.7 kB/s	0 B/s
PACKETS_RECEIVED (r1 : Ethernet10)	0 p/s	1.46 kp/s	2.64 kp/s	0 p/s
PACKETS_TRANSMITTED (r1 : Ethernet10)	0 p/s	728 p/s	1.33 kp/s	0 p/s
BYTES_RECEIVED (r2 : Ethernet1) (right y-axis)	0 B/s	2.22 MB/s	4.04 MB/s	0 B/s
BYTES_TRANSMITTED (r2 : Ethernet1) (right y-axis)	0 B/s	51.1 kB/s	94.1 kB/s	0 B/s
PACKETS_RECEIVED (r2 : Ethernet1)	0 p/s	1.46 kp/s	2.66 kp/s	0 p/s
PACKETS_TRANSMITTED (r2 : Ethernet1)	0 p/s	727 p/s	1.34 kp/s	0 p/s
BYTES_RECEIVED (r2 : Ethernet10) (right y-axis)	0 B/s	51.1 kB/s	94.1 kB/s	0 B/s
BYTES_TRANSMITTED (r2 : Ethernet10) (right y-axis)	0 B/s	2.22 MB/s	4.04 MB/s	0 B/s
PACKETS_RECEIVED (r2 : Ethernet10)	0 p/s	727 p/s	1.34 kp/s	0 p/s
PACKETS_TRANSMITTED (r2 : Ethernet10)	0 p/s	1.46 kp/s	2.66 kp/s	0 p/s

```
[  5] 292.00-293.00 sec  3.75 MBytes  31.5 Mbits/sec    0    225 KBytes
[  5] 293.00-294.00 sec  3.75 MBytes  31.5 Mbits/sec    5    170 KBytes
[  5] 294.00-295.00 sec  3.75 MBytes  31.5 Mbits/sec    0    204 KBytes
[  5] 295.00-296.00 sec  2.50 MBytes  21.0 Mbits/sec    0    222 KBytes
[  5] 296.00-297.00 sec  3.75 MBytes  31.5 Mbits/sec    2    171 KBytes
[  5] 297.00-298.00 sec  3.75 MBytes  31.5 Mbits/sec    0    185 KBytes
[  5] 298.00-299.00 sec  2.50 MBytes  21.0 Mbits/sec    0    192 KBytes
[  5] 299.00-300.00 sec  3.75 MBytes  31.5 Mbits/sec    0    204 KBytes
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate         Retr
[  5]   0.00-300.00 sec  1013 MBytes  28.3 Mbits/sec  258             sender
[  5]   0.00-300.06 sec  1009 MBytes  28.2 Mbits/sec                  receiver

iperf Done.
```

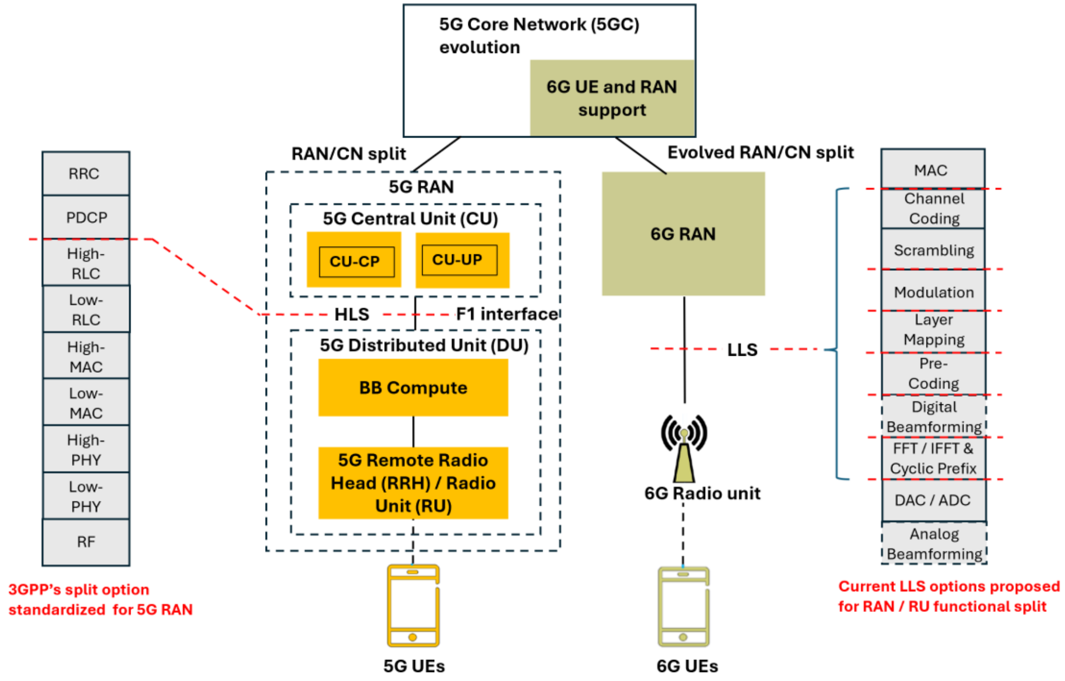
```
[  5] 294.00-295.00 sec  4.25 MBytes  35.7 Mbits/sec
[  5] 295.00-296.00 sec  3.06 MBytes  25.7 Mbits/sec
[  5] 296.00-297.00 sec  3.36 MBytes  28.2 Mbits/sec
[  5] 297.00-298.00 sec  3.03 MBytes  25.4 Mbits/sec
[  5] 298.00-299.00 sec  3.24 MBytes  27.2 Mbits/sec
[  5] 299.00-300.00 sec  2.86 MBytes  24.0 Mbits/sec
[  5] 300.00-300.06 sec   204 KBytes  29.8 Mbits/sec
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bitrate
[  5]   0.00-300.06 sec  1009 MBytes  28.2 Mbits/sec
-----------------------------------------------------------
Server listening on 5201
-----------------------------------------------------------
^Ciperf3: interrupt - the server has terminated
```

Results on E2E system evaluation by simulations

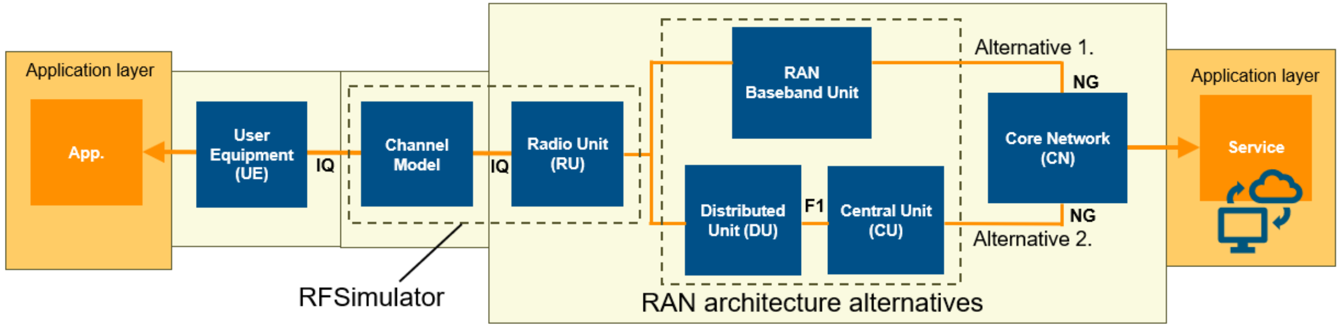
Simulations on 6G RAN – Simulation scenario

- A simulation study has been done for the latency evaluation of 5G disaggregated RAN and its comparison to monolithic RAN architecture.

5G and 6G RAN approaches & HLS / LLS options.



Studied RAN split alternatives simulator high-level architecture.



Simulation environment characteristics & radio configuration parameters.

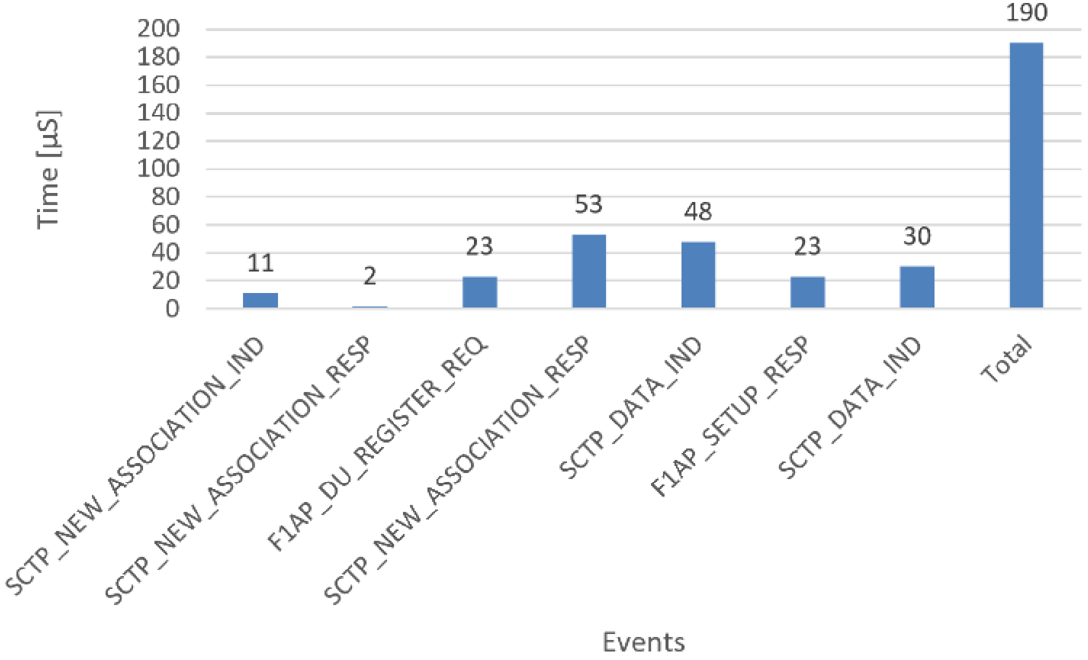
Hardware	Intel i7-11850H @2.5G, 16 Core, 32 GB RAM
Operating System	Ubuntu 20.0.6 LTS
Number of TX antennas	2
Number of RX antennas	2
Frequency Band	n78, ~3500MHz
SCS	30 kHz
Bandwidth	40 MHz
Duplexing	TDD
Channel Model	AWGN
Simulation Mode	Standalone NR 5G

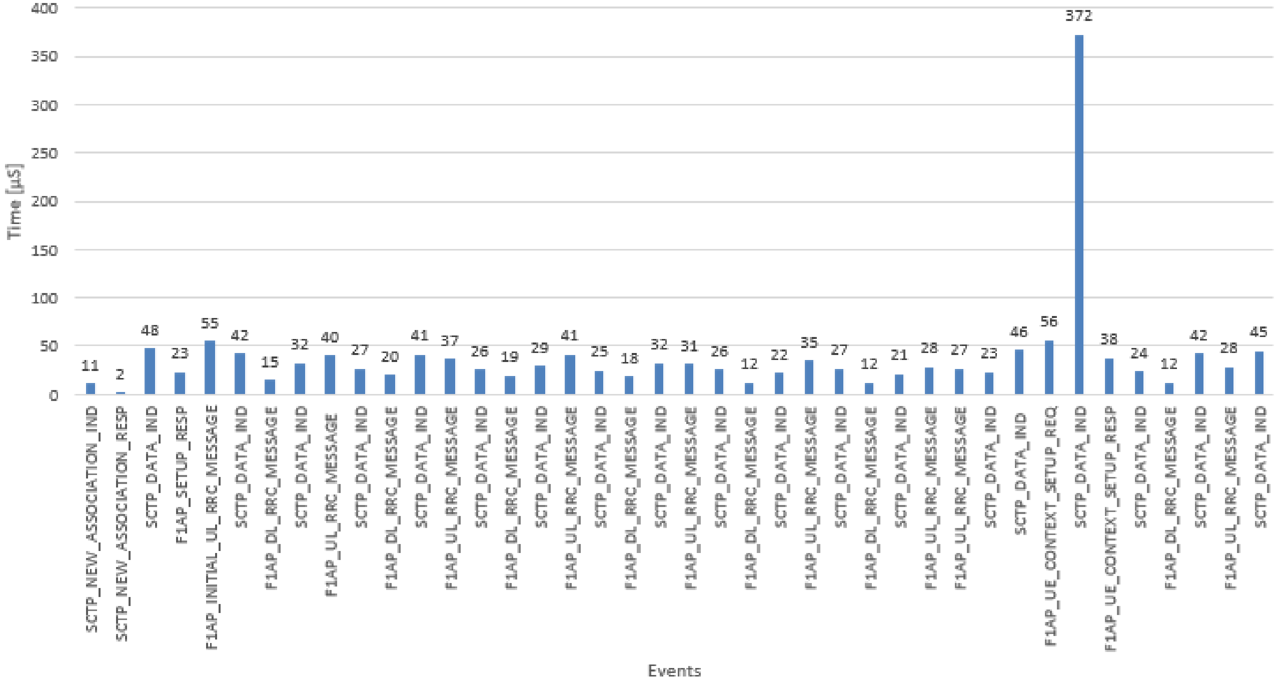
Simulations on 6G RAN – Simulation results (1)

Disaggregated RAN case (simulation architecture Alternative 2)

- The procedure for disaggregated RAN is as follows:
  - ○ Start CU process
  - ○ Start DU process which triggers F1 startup
  - ○ Start UE process for initial access
  - ○ Collect measured processing times of the message exchange.

F1 interface establishment between DU & CU events time durations.





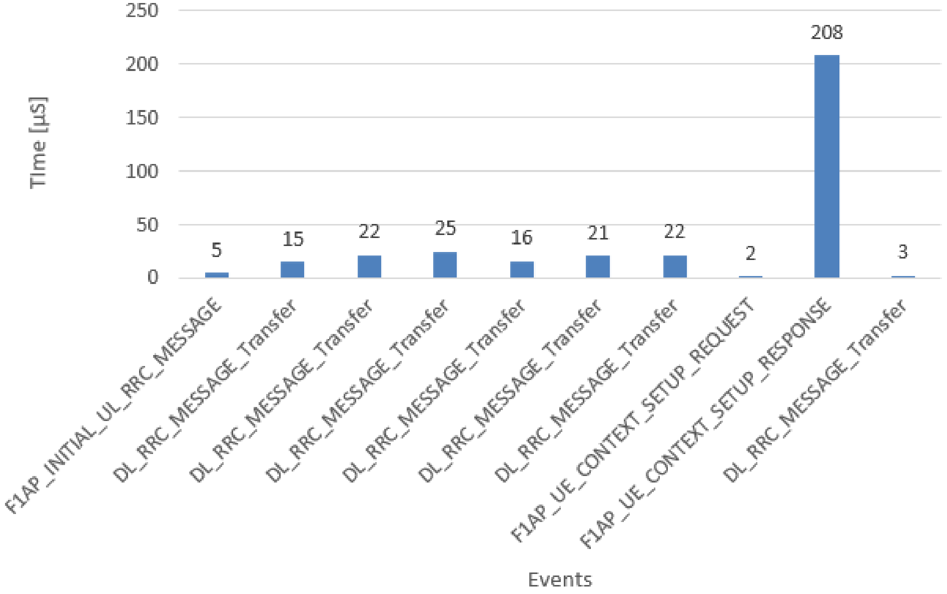
Time durations of events at DU & CU during UE initial access procedure.

Simulations on 6G RAN – Simulation results (2)

Monolithic RAN case (simulation architecture Alternative 1)

- The simulation procedure for monolithic RAN is as follows:
  - ○ Start monolithic gNB process
  - ○ Start UE process for initial access
  - ○ Collect measured processing times of the message exchange.

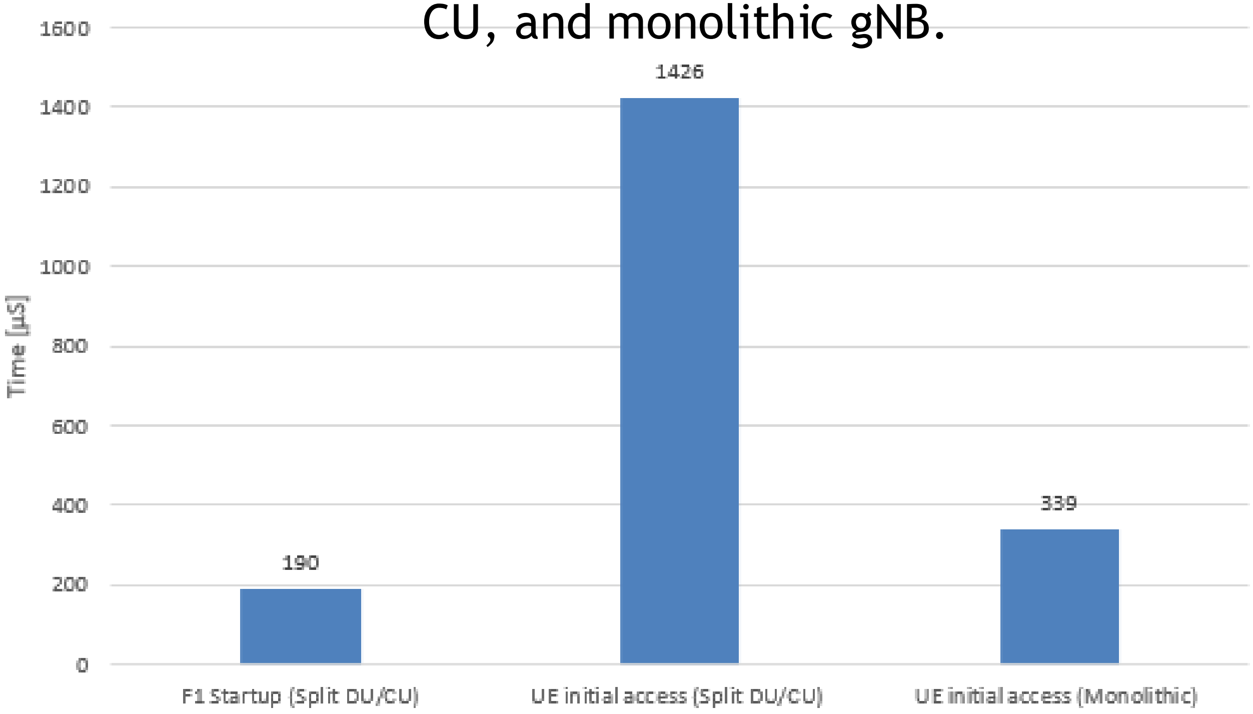
UE initial access procedure in monolithic RAN case time durations events.



Discussion:

- Monolithic case has substantially lower latency for UE initial access process.
  - ○ No need of SCTP_DATA_IND messages.
- Physical distance would increase delay in the disaggregated RAN case.

Simulated total times for split DU & CU, and monolithic gNB.



Chapter 4
Evaluation of the security, privacy and system-level resilience

Addressing the goal of a trustworthy environment and supporting resilience and availability

- As expressed by the Hexa-X-II design principles 5 (resilience and availability) and 6 (security and privacy)
  - *Establish a comprehensive framework in the 6G E2E system that ensures security and privacy are integrated across all components with the goal of assuring a trustworthy environment. E.g., address current as well as future threats in a resilient manner against attack and incorporate security fundamentals in its design, inherently support the preservation of privacy, and allow different levels of anonymity for future services.(Principle 6)*
  - *The network should allow mobile network operators (MNOs) to build deployments with high resilience and availability. (Principle 5)*
- Structured according to the TRL of each of the security controls being considered:
  - Basic TRLs (TRL 1 and 2):

    →Fundamentally related to Physical Layer Security (PLS).

    →Conceptual evaluation, by simulation or experimental laboratory environments.
  - Validation TRLs (TRL 3 and 4):

    →Trustworthy AI, quantum-resistant crypto and simulated resilience evaluation.

    →Experimental measurements and proposal for integration patterns.
  - Development TRLs (TRL 5 and 6):

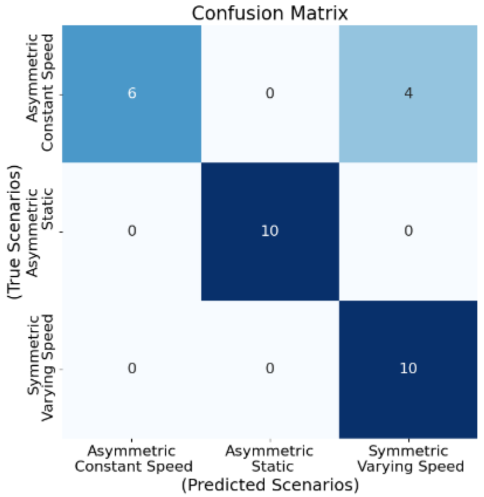
    →Confidential network deployment

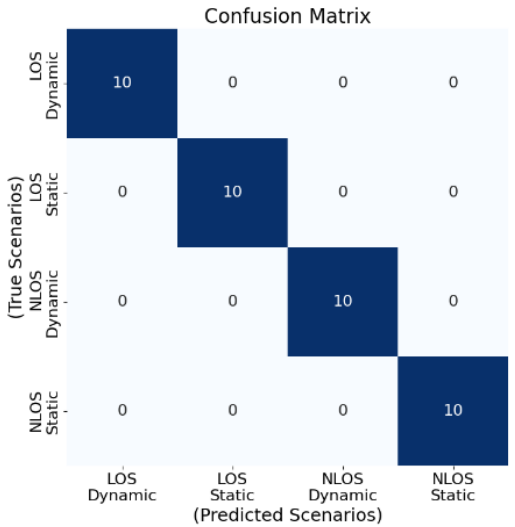
    →Distributed ledgers.

Enablers at Basic TRLs

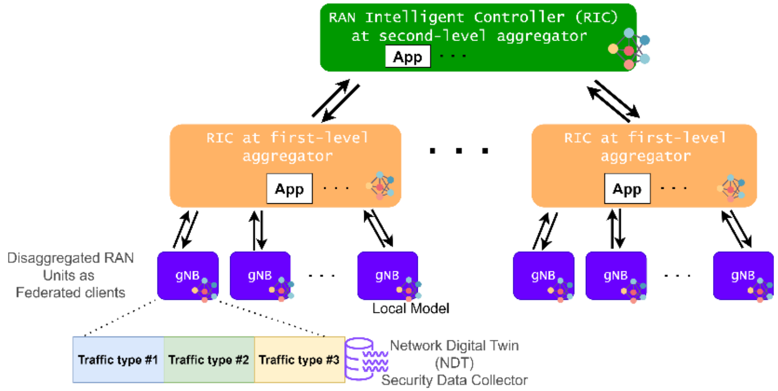
Physical Layer Security

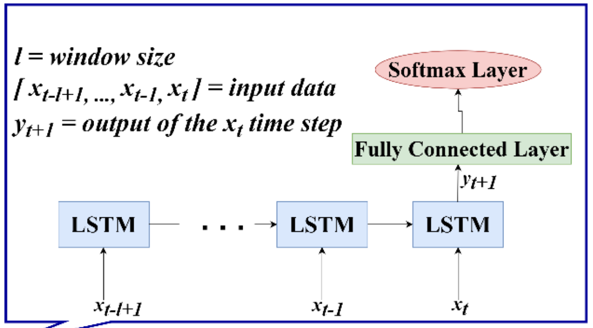
- Physical context awareness:
    - Adapt security controls to environmental conditions.
    - Considering graceful degradation.
    - Recognize the condition of the wireless channel to adjust key generation.
    - Experiments with two implementations highlighting relevant physical parameters.



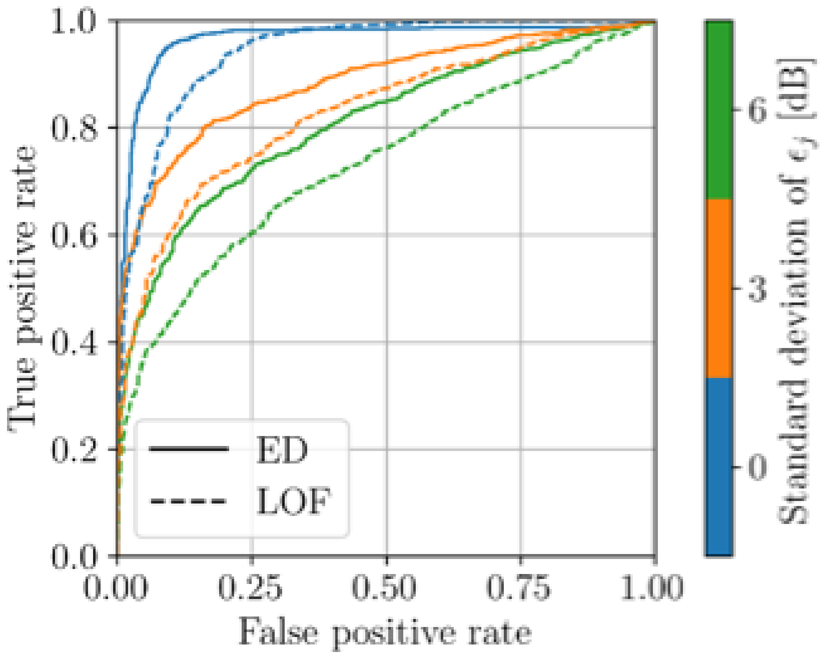


- Physical anomaly detection:
    - Detect PHY layer anomalies.
    - Discriminate casual interference from malicious jamming.
    - Relying on the use of NDT techniques.


RAN Intelligent Controller (RIC) at second-level aggregator
App
RIC at first-level aggregator
App
RIC at first-level aggregator
App
Disaggregated RAN Units as Federated clients
gNB
gNB
gNB
gNB
gNB
gNB
Local Model
Traffic type #1
Traffic type #2
Traffic type #3
Network Digital Twin (NDT) Security Data Collector

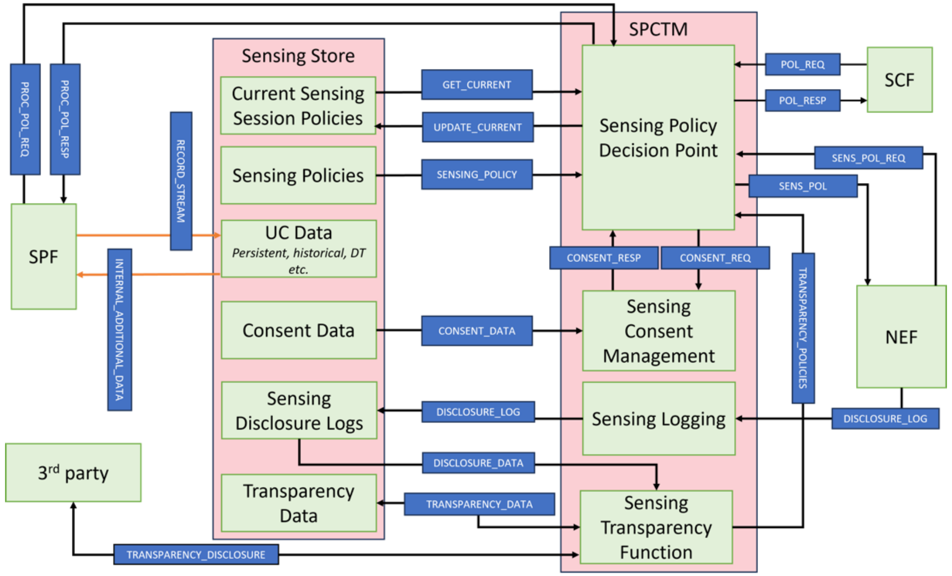


LSTM classification model architecture

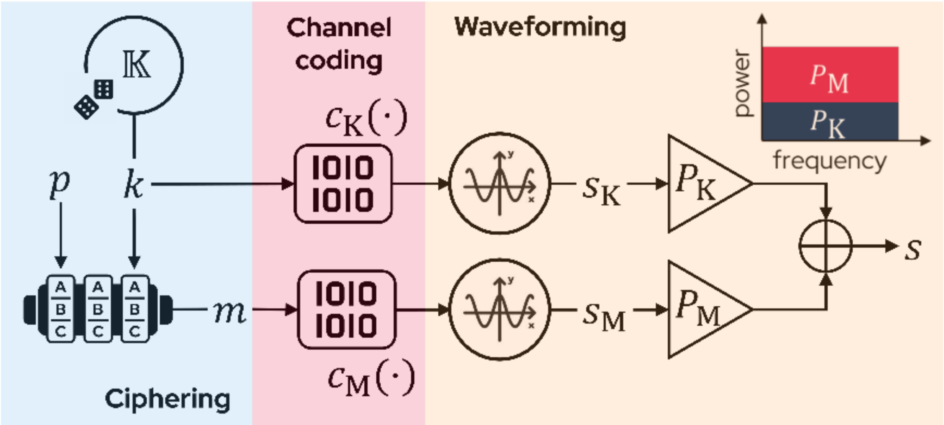


Physical Layer Security

- JCAS threat mitigation:
  - Authentication and access control to sensing information flows.
  - Confidentiality and integrity for sensing data in transit and at rest.
  - Sensing operation as per the consent information from UE and targets .
  - Data authentication, and secure export and import of data.
  - Safeguarding synchronization mechanisms on the radio side.



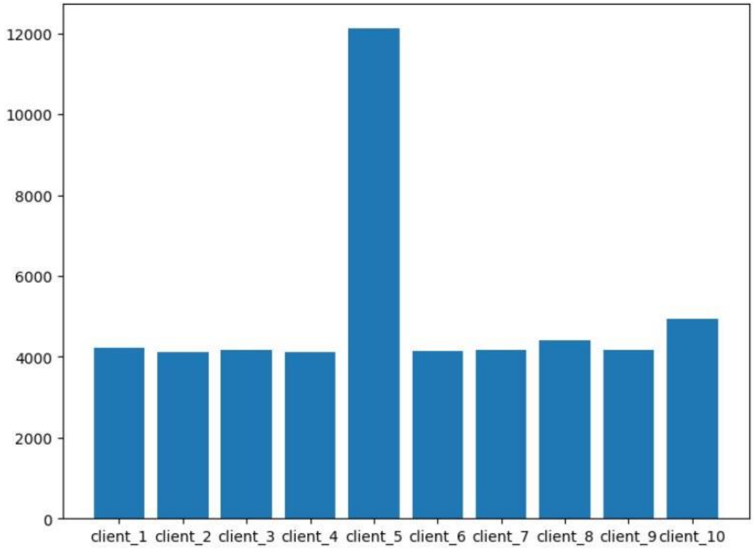
- Physical layer deception:
  - Deceive potential eavesdroppers.
  - Transmit compromised information while simultaneously delivering the original data to authorized receivers.
  - Optimized regarding the channel coding rate of and the power multiplexing ratio.



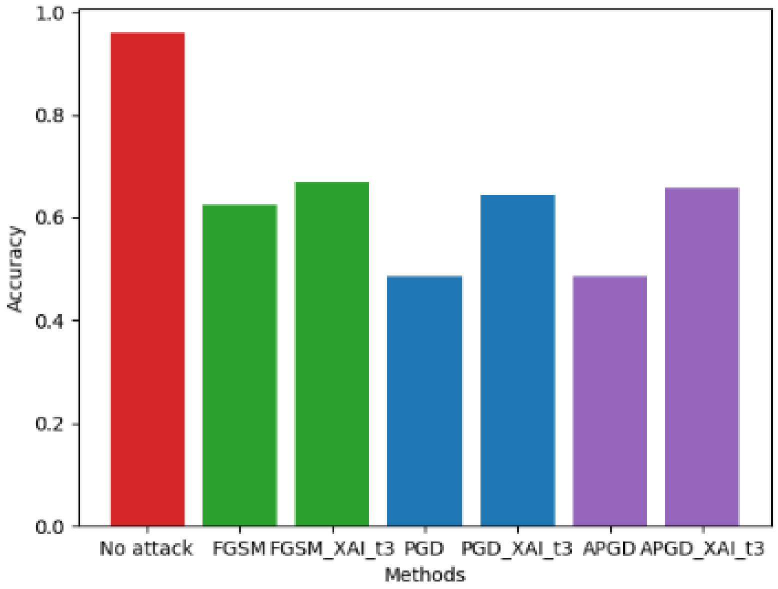
Enablers at Validation TRLs

Trustworthy AI

- A framework for security and privacy in federated learning:
  - Allow the usage of secure aggregation schemes.
  - Evaluation of the effectiveness in preventing security and privacy attacks:
    - Malicious client identification.
    - Gradient hiding mechanisms to increase privacy protection level.
  - Analyze computation and communication overhead.



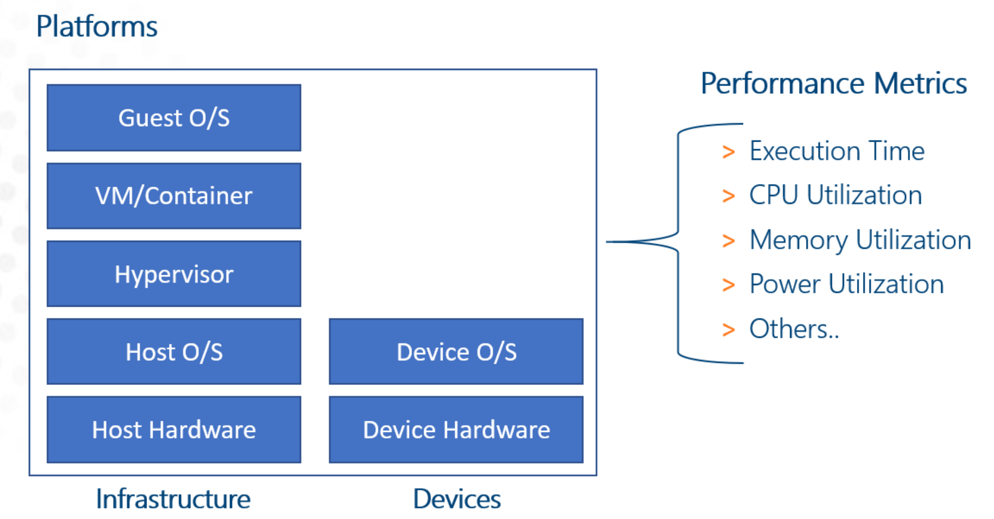
- Explainable AI against adversarial attacks on IDS:
  - Crafted to deceive AI-enabled IDS to misclassify malicious DDoS traffic as normal.
  - Transparency allows for a deeper understanding of the behavior of the AI model.
  - Apply it to autonomously fine-tune for improved accuracy and performance.
  - More accurate and efficient identification of anomalous network behavior.

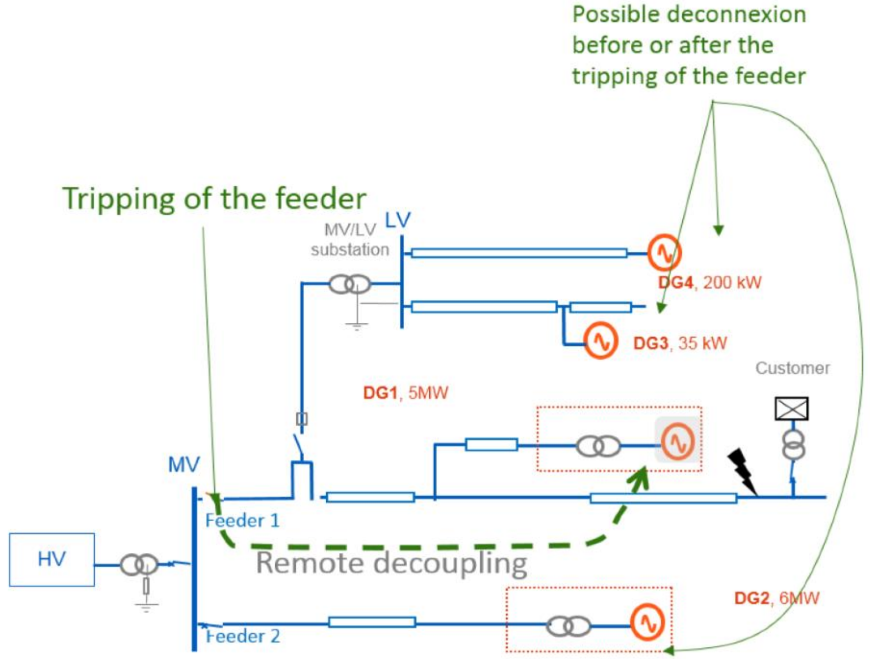


Cryptography and Resilience

- PQC evaluation based on the *Qujata* platform:
  - Explore additional algorithms.
  - Analyze hybridization approaches, including strategies for QKD integration.
  - Consider additional metrics to evaluate the impact on secure protocols:
    - →Induced latency.
    - →Deviations in execution time that can translate into jitter.

- E2E resilience simulations:
  - Two kinds of operationally relevant simulations.
  - VNF availability, considering the Impact of self-healing in microservice environments.
  - Resilience studies guided by vertical requirements, with an exemplary case on the electrical grid.

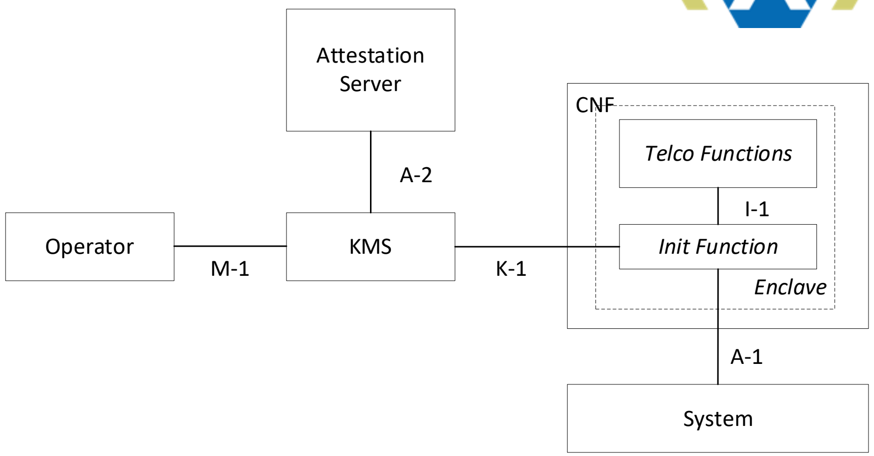


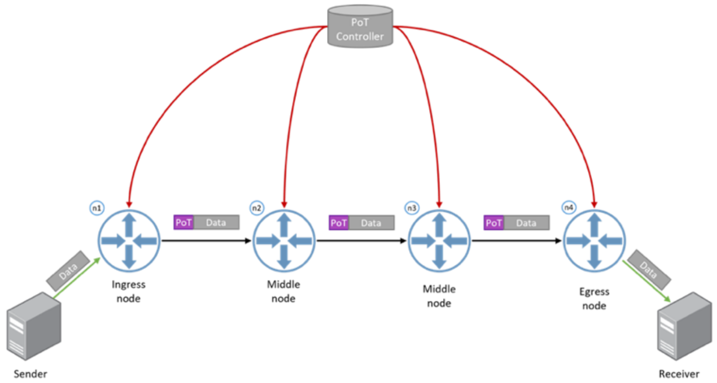


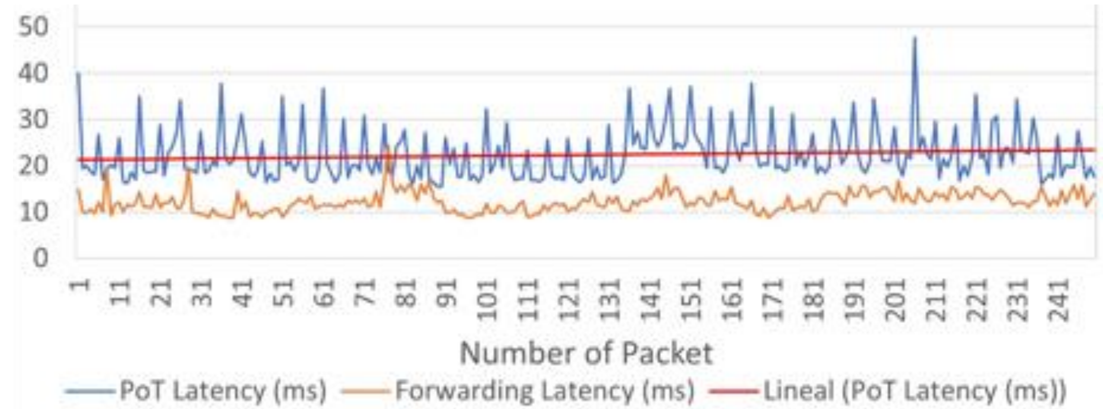
Enablers at Development TRLs

Confidential Network Deployment

- Relevant aspects of the application of Confidential Computing:
  - Performance
    - → Focused on the use of secure enclaves.
    - → Performance impact of enclave paging and cache misses.
    - → Evolve the design of NFs or even network architecture to facilitate data locality.
  - Operations and standardization
    - → The relevance of proper key management.
    - → Standards for key management and remote attestation.
- Impact on network performance of topology attestation:
  - Measuring latency and throughput in PoT-enabled NF chains.
  - Disavow continuous attestation, but focus on IOAM practices.
  - Linear dependency on chain length.
  - Operational recommendations for IOAM.





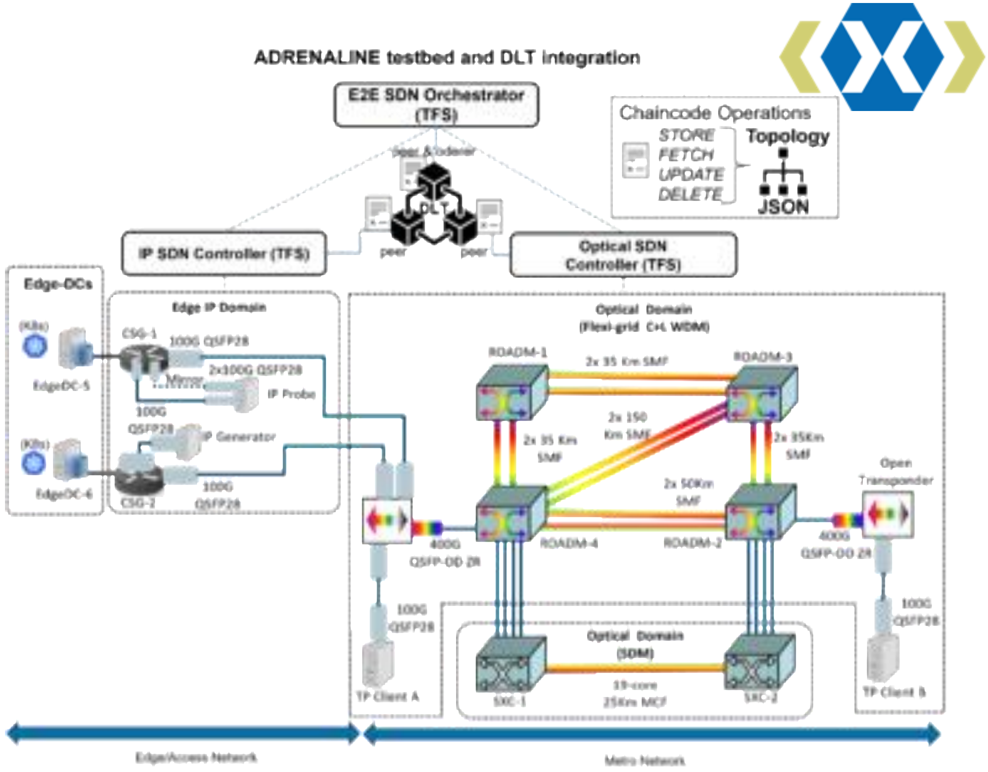


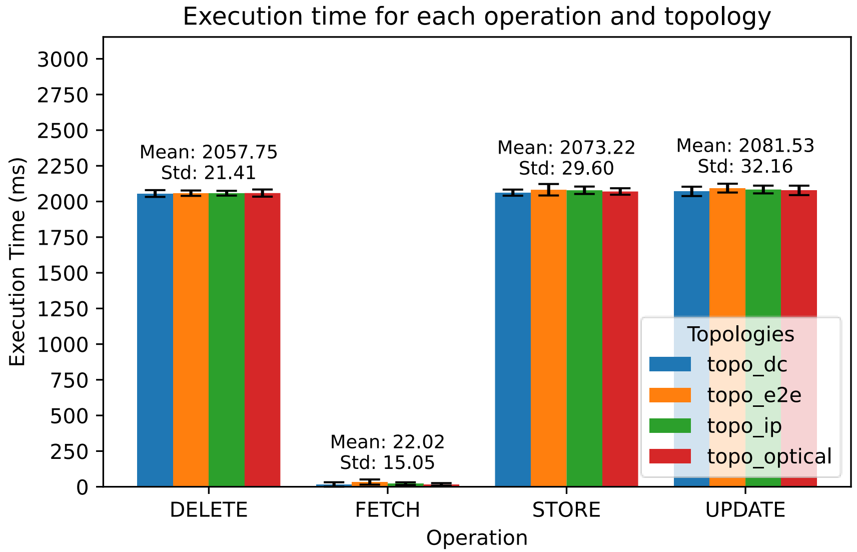
Distributed Ledgers

- Demonstrate DLTs with a permissioned ledger approach to topology management:
  - Changes recorded as transactions.
  - Enhance security, transparency, and trustworthiness across multiple domains.
  - Cases of connectivity service orchestration combining domain SDN controllers.

- Evaluate the performance of DLT operations as support for orchestration:
  - Higher execution times are those that write/change the state of the ledger.
  - No significant impact when the topology description size.

- Prototype a DLT gateway for events related to connectivity services:
  - API methods and typedefs.





Chapter 5 Conclusions

System design validation(1)

Design Principle	Evaluated Aspects	Involved System-PoC Components	Proof Points (highlights)
1: Support and exposure of 6G services and capabilities	Compute service provision (offloading), intent-based service exposure, in-network service exposure	Service autoscaling and migration, trustworthy flexible topologies, intent-based network solution, management capabilities exposure framework	**Compute offloading successfully demonstrated** for cobots, **AI-based lifecycle management of 6G latency-sensitive** across the cloud continuum. **Management capabilities exposure** framework provides a communication channel and exposes capabilities to other (third-party) entities. **Intent-based management with NLP** processing of user requests.
2: Full automation and optimization	Pervasive infrastructure for pervasive service orchestration, multi-platform orchestration, closed-loop control and distributed AI/ML agents for automation and optimization	Service and resource orchestration, closed-loop automation, AI-assisted E2E life-cycle management (service autoscaling and migration)	**Zero-touch** cobot-based video surveillance **with NLP processing and service instance creation within 250s**. Resource orchestration demonstrates functionality to **deploy services across the cloud-continuum provided by different providers. Closed-loop automation demonstrated** for cobot service migration. **RL-based joint autoscaling and placement/migration for 6G latency-sensitive** services meet E2E latency requirements.
3: Flexibility to different network scenarios	Pervasive service management and orchestration with multi-cluster resource orchestration, federated orchestration, flexible topologies	Service and resource provisioning, DLT-based service federation, trustworthy flexible topologies and beyond communication aspects, Programmable and flexible network configuration	**Multi-cluster resource orchestration** brings flexibility in service deployment. **Federated orchestration** allows dynamic connections to third-party networks. **Flexible topologies** adapt to traffic bursts and user requirements. Programmable and flexible network configuration with SDN controller **seamlessly integrate and manage both the packet and optical layers of transport networks.**
4: Scalability	Pervasive service management and orchestration, decentralized orchestration components to adapt to traffic and consumer needs	Service autoscaling and migration, extreme-edge cluster emulation, DLT-based service federation	RL-based joint autoscaling and placement/migration for 6G latency-sensitive services demonstrated. **DLT-based service federation enables resource sharing between providers** when additional resources are needed.

System design validation (2)

Design Principle	Evaluated Aspects	Involved System-PoC Components	Proof Points (highlights)
5: Resilience and availability	Pervasive service management and orchestration to guarantee service continuity, flexible network topologies improving availability and reliability of the network; E2E resilient network dimensioning	Service autoscaling and migration, ;closed-loop automation, trustworthy flexible topologies, E2E resilience planning tool	**High availability** demonstrated in an emulated extreme-edge infrastructure for real-time video streaming with zero-touch cobobot-based videosurveillance. E2E resilience simulations demonstrate the ability to achieve high availability (above 5 nines) for critical use cases like remote power station protection.
6: Persistent security and privacy	Comprehensive framework for security and privacy integration	physical layer security, trustworthy AI, quantum resistant cryptography, level of trust assessment, confidential network deployment, distributed ledgers	Evaluation results provide **evidence on the feasibility of enablers and SPR controls**. Results are provided according to planned TRL for each enabler.
7: Internal interfaces are cloud optimized	Cloud-native approach to service development and orchestration	Inter-cluster communication, management capabilities exposure framework	Two alternative options for **inter-cluster communication** explored: karmada multi-cluster management tool and Network Service Mesh. **Functional validation for the management capabilities exposure** framework reported.
10: Minimize environmental footprint and enabling sustainable use cases	E2E orchestration with energy-aware placement algorithms, energy-efficient offloading	Functionality Allocation function in trustworthy flexible topologies	**Functionality Allocation** function in trustworthy **flexible topologies** provides **energy consumption gains** in cobot-powered warehouse inventory management scenario.

Proof points highlight underscore key results that validate adherence to the design principles

Conclusion and next steps

- **System Design Validation**
  - The current system design for 6G (i.e., Hexa-X-II blueprint) emphasizes several key principles (e.g., full automation, scalability, resilience, and security) by integrating different solutions for intent-based network, closed-loop automation, and AI-based scenarios.
  - The system can support a wide range of network scenarios, optimize resource usage, and maintain high levels of performance and availability.
  - The current system design principles set forth in D2.1 are effectively met, providing a robust foundation for future 6G developments.

- **Performance and Key Value Indicators**
  - System-PoC #B's evaluation results highlight the system's ability to meet the stringent KPIs and KVIs essential for 6G networks.
  - Focus on latency-sensitive services (i.e., cobots real-time video surveillance) underscores the system's capability to handle compute-intensive operations with ultra-low latency.
  - Intent-based service provisioning and closed-loop automation ensures zero-downtime and optimal resource utilization.
  - The incorporation of energy-efficient operation policies, resilient network topologies, and reduced operational costs aligns with the sustainability goals.

Conclusion and next steps

- **Security and Privacy**
  - The integration of comprehensive security frameworks (e.g., quantum-resistant cryptography, trustworthy AI, confidential network deployment, etc.) ensures a secure and trustworthy environment for 6G services.
  - The project's focus on addressing current and future threats through inherent design principles and advanced security measures signifies a proactive approach to safeguarding 6G networks.

- **Future Directions**
  - The insights gained from System-PoC #B will allow a continuous improvement and refinement of the 6G system designs, ensuring that future iterations are well-equipped to meet the evolving demands of next-generation communication networks.
  - This deliverable will trigger the last phase (i.e., System-PoC #C) with new results from the current enablers and solutions and with the addition of other technologies such as radio communications.

HEXA-X-II

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.