Call: H2020-ICT-2020-2

Project reference: 101015956

Project Name:

A flagship for B5G/6G vision and intelligent fabric of technology enablers connecting human, physical, and digital worlds

Hexa-X

Deliverable D7.2
Special-purpose functionalities: intermediate solutions

Date of delivery:	30/04/2022	Version:	1.0
Start date of project:	01/01/2021	Duration:	30 months

Document properties:

Document Number:	D7.2
Document Title:	Special-purpose functionalities: intermediate solutions
Editor(s):	Bin Han (TUK), Bjoern Richerzhagen (SAG), Lucas Scheuvens (TUD), Claudio Casetti (POL)
Authors:	Bin Han (TUK), Bjoern Richerzhagen (SAG), Lucas Scheuvens (TUD), Claudio Casetti (POL), Alberto Martinez Alba (SAG), Ioannis Belikaidis (WIN), Serge Bories (CEA), Carla-Fabiana Chiasserini (POL), Claudio Demartini (POL), Panagiotis Demestichas (WIN), Valentina Gatteschi (POL), Mohammad Asif Habibi (TUK), Dinh-Thuy Phan Huy (ORA), Antonis Karaolanis (WIN), Ingolf Karls (INT), Thodoris Kasidakis (WIN), Fabrizio Lamberti (POL), Vasiliki Lamprousi (WIN), Mattia Merluzzi (CEA), Honglei Miao (INT), Martti Moisio (NOF), Riccardo Rusca (POL), Hans Dieter Schotten (TUK), Mohammad Shehab (OUL), Emilio Calvanese Strinati (CEA), Eduardo Tominaga (OUL), Karthik Upadhya (NOF), Amir Varastehhajipour (SAG)
Contractual Date of Delivery:	30/04/2022
Dissemination level:	PU[1]
Status:	Final
Version:	1.0
File Name:	Hexa-X D7.2_v1.0

Revision History

Revision	Date	Issued by	Description
	16.12.2021	WP7	Initial Table of Contents / structure
	07.02.2022	WP7	Initial draft of technical content from individual tasks
	09.03.2022	WP7	Version for cross-task reviews
	17.03.2022	WP7	Complete version for project internal review
	01.04.2022	WP7	Complete version for GA review
	28.04.2022	WP7	Final version for submission to EC

Abstract

This report presents the advancements in scope of Working Package 7 (WP7) regarding the identified special-purpose functionalities, including new concepts, proposed solutions, intermediate analyses and demonstration results.

[1] CO = Confidential, only members of the consortium (including the Commission Services)

PU = Public

Keywords

Dependability, sustainable coverage, I4.0, radio resource management, digital twin, key performance indicators

Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101015956.

Executive Summary

This report “Special-purpose functionalities: intermediate solutions” is the second deliverable of Work Package 7 (WP7) and focuses on technical enablers for extreme experiences in Internet-of-Things and Industry 4.0 environments. The deliverable contains technical proposals to close the gaps identified in Deliverable 7.1 – “Gap analysis and technical work plan for special-purpose functionality”, and their intermediate results. This report updates the use cases and key performance/value indicators that are introduced in Deliverable 1.2 – “Expanded 6G vision, use cases and societal values – including aspects of sustainability, security and spectrum” – and thereafter discussed in Deliverable 7.1.

Based on the updated use cases and KPIs/KVIs, three technical tasks are identified to overcome the challenges of extreme experiences: ultra-flexible resource allocation caused by the bandwidth crunch in traditional radio bands, modelling the 6G dependability in future I4.0 environments, and massive deployment of Digital Twin (DT) in combination with novel human-machine interface (HMI).

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Term	Description
3GPP	Third Generation Partnership Project
4G	Fourth Generation of Wireless Communications Systems
5G	Fifth Generation of Wireless Communications Systems
5G-PPP	5G Infrastructure Public Private Partnership
6G	Sixth Generation of Wireless Communications Systems
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
AR	Augmented Reality
B5G	Beyond 5G
BS	Base Station
CCDF	Complementary Cumulative Distribution Function
CD	Crowd-Detectable
CDF	Cumulative Distribution Function
CF	Cell-Free
CR	Collaborative Robot
CU	Centralized Unit
cMTC	Critical Machine-Type Communications
CNN	Convolutional Neural Network
CPE	Cyber-Physical Environment
CPU	Central Processing Unit
DT	Digital Twin
DU	Distributed Unit
E2E	End-to-End
EC	European Commission
ECG	Electrocardiogram
EEG	Electroencephalogram
EI	Emergent Intelligence
EMF	Electromagnetic Field
EMG	Electromyogram
EOG	Electrooculogram
ES	Edge Server
EU	European Union
FA	Functionality Allocation

FE	Functional Entity
FL	Federated Learning
GPS	Global Positioning System
H2020	Horizon 2020
HARQ	Hybrid Automatic Repeat Request
HE	Hosting Entity
HMD	Head-Mounted Display
HMI	Human-Machine Interface
HTC	Human-Type Communications
HW	Hardware
I4.0	Industry 4.0
ICT	Information and Communication Technologies
In^2-X	Intra-X and Inter-X
IoT	Internet of Things
KPI	Key Performance Indicator
KVI	Key Value Indicator
LTI	Linear-Time-Invariant
LTV	Linear-Time-Variant
LU	Local User
MAC	Medium Access Control
MEC	Multi-Access Edge Computing
MIMO	Multiple-Input Multiple-Output
MJLS	Markov Jump Linear System
MKS	Mouse-Keyboard-Screen
ML	Machine Learning
mMIMO	Massive MIMO
MR	Mixed Reality
MTC	Machine-Type Communications
MTTF	Mean Time till Failure
NACK	Negative Acknowledgement
NBI	North-Bound-Interface (NBI)
NLOS	Non-Line-of-Sight
NPN	non-public or private networks (NPN)
OSM	Open-Source Management and Orchestration
PD	Partially Distributed
PDU	Packet Data Protocol Unit
PER	Packet Error Rate
PHY	Physical
PLR	Packet Loss Rate

QoS	Quality of Service
RAN	Radio Access Network
REM	Radio Environment Map
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RU	Remote User
SIM	Subscriber Identification Module
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Radio
SW	Software
TCO	Total Cost of Ownership
TCP	Transmission Control Protocol
TD	Totally Distributed
WP	Work Package
UAV	Unmanned Aerial Vehicle
UE	User Equipment
UI	User Interface
URLLC	Ultra-Reliable Low-Latency Communications
VR	Virtual Reality
X2I	X-to-Infrastructure
XR	Extended Reality
ZED	Zero-Energy-Device

1 Introduction

Hexa-X is one of the 5G-PPP projects under the European Union (EU) Horizon 2020 framework. It is a flagship project that develops a Beyond 5G (B5G)/6G vision and an intelligent fabric of technology enablers connecting human, physical and digital worlds.

This report is the second deliverable of Work Package 7 (WP7) – "Special-Purpose Functionality". It presents a deepened analysis of selected use cases defined by Work Package 1 in Deliverable D1.2 "Expanded 6G vision, use cases and societal values" [Hexa-X-D1.2], with a focus on dependability and spatial coverage in Industry 4.0 (I4.0) / Internet of Things (IoT) environments. It proposes technical solutions to close the gaps that are identified and analysed in D7.1 [Hexa-X-D7.1], with intermediate results. The technical contributions delivered in this document align with the work plan for Hexa-X WP7 that is formulated in D7.1.

1.1 Objective of the document

The objective of this document is to address the technical challenges identified in D7.1 and the project objectives defined in [Hexa-X-D1.2] in scope of I4.0/IoT scenarios. Technical solutions are included to fulfill the B5G/6G service performance requirements refined for such scenarios and to close the research gaps regarding the planned contributions in the project.

This document guides the work in WP7 and provides input to other Hexa-X work packages from a technical perspective.

1.2 Structure of the document

This document is structured as follows. Chapter 2 provides an assembled overview to the contributions of WP7 in addressing the Hexa-X uses cases and objectives, and their relations to the end-to-end 6G architecture as well as to other Hexa-X technical enablers. Chapter 3 reports the advancements in the use cases investigated in WP7 and dealing with the need for ultra-flexible resource allocation caused by the bandwidth crunch in traditional radio bands. Chapter 4 reports the advancements in the modelling, evaluation, and applications of the 6G dependability concept in future I4.0 environments. Chapter 5 introduces the proposed new concepts, use cases, and intermediate technical solutions regarding massive deployment of Digital Twin (DT) in combination with novel human-machine interface (HMI). To the end, Chapter 6 concludes this document.

2 Overview of initial solutions and their relation to use cases, architecture, and technical enablers

This section provides an overview of the special-purpose functionality being proposed to increase B5G/6G system performance in extreme environments. We map initial solutions to the targeted use cases and their Key Performance Indicators (KPIs) / Key Value Indicators (KVIs) based on recent updates in [Hexa-X D1.3] and discuss the embedding in the initial end-to-end architecture based on the architectural goals and enablers presented in [Hexa-X D5.1]. Lastly, we highlight relations to other technical enablers from within the Hexa-X project that are utilized in close collaboration to realize the envisioned solutions for special-purpose functionality. As discussed in [Hexa-X D7.1], our work focuses on two groups of use cases from [Hexa-X D1.3]: use cases related to *Dependability in Industry 4.0* scenarios, and use cases related to *Sustainable Coverage in IoT*, as shown in Figure 2-1.

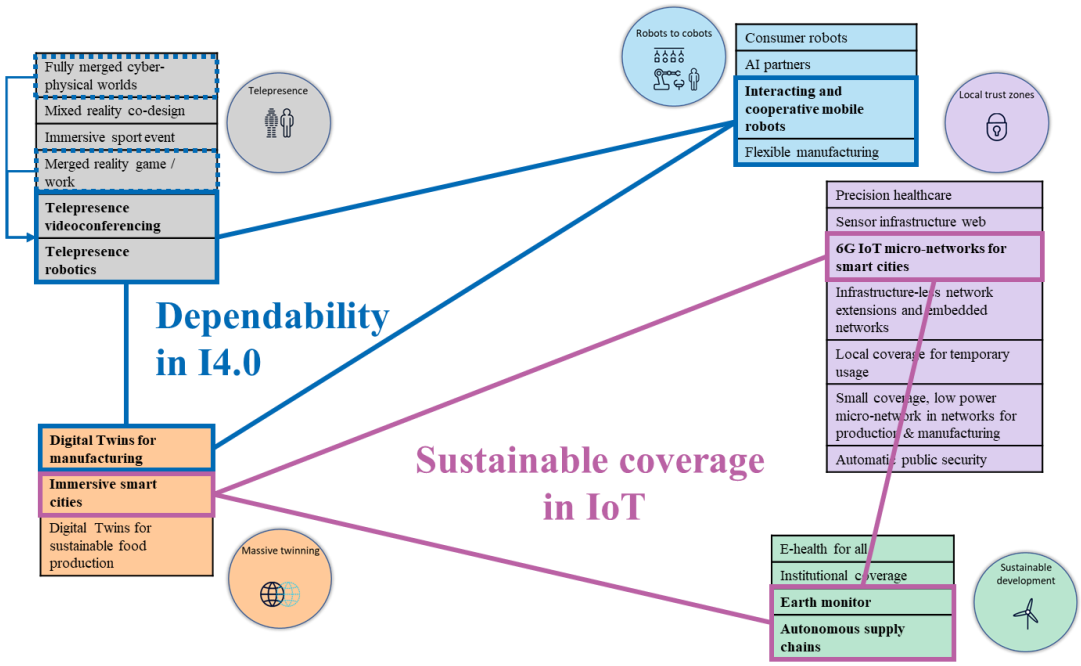


Figure 2-1: Use cases grouped into Dependability in I4.0 and Sustainable Coverage in IoT [Hexa-X D7.1].

2.1 Overview of initial solutions

To fulfil the requirements for the use cases outlined in [Hexa-X D7.1], three objectives are defined for special-purpose functionality, as summarized in Table 1. Our work towards Objective 1 explicitly targets both groups of use cases: dependability in I4.0 and sustainable coverage in IoT. Contributions to Objectives 2 and 3 focus on use cases related to dependability in I4.0.

Table 1: Objectives for special-purpose functionality.

Objective	Description	Targeted use cases
Obj. 1	Define *ultra-flexible resource allocation procedures* in challenging environments such as those populated by mobile devices with special requirements and in need of coverage.	Dependability in I4.0, sustainable coverage in IoT
Obj. 2	Develop mechanisms and enablers for *high dependability in vertical scenarios*, enabling efficient resource support of complex and dynamically changing availability requirements.	Dependability in I4.0
Obj. 3	Support the convergence of the biological, digital and physical worlds with human interaction through *novel HMI concepts and privacy-preserving high-availability Digital Twins*.	Dependability in I4.0

Objectives 1 and 2 require an understanding of the *resources* and *mechanisms* that can be utilized in the respective environment to increase performance. This includes identifying relevant *environmental conditions* and resulting *influences* that need to be captured and modelled as a basis for decision making in resource allocation, both for sustainable coverage and increased dependability as illustrated in Figure 2-2. The model on the right-hand side of the figure is referred to as the Digital Twin, which is then utilized in decision making and resource allocation, but also for novel human-machine interaction. While depicted as a single entity in the figure, we propose a distributed, collaborative approach to Digital Twins addressing Objective 3, as later discussed in Section 5.

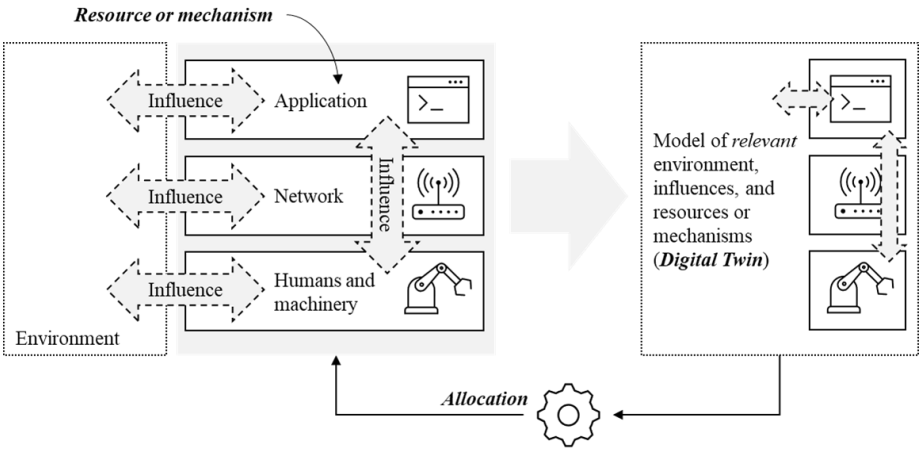


Figure 2-2: Concept of influences and their model in a digital twin used, e.g., for resource allocation.

The influence relations identified and addressed in our work are summarized in Table 2, with pointers to the respective contributions in terms of resources or mechanisms discussed in this deliverable. Influences are bi-directional, i.e., entries influence both, the column and row they are located in. In the table, *network* also refers to novel 6G capabilities such as Compute- or AI-as-a-Service. Underlined contributions are part of a demonstrator, with core aspects being discussed in Section 4.2.

Table 2: Bi-directional influences relations and respective resources/mechanisms discussed in this deliverable, with underlined entries being part of a demonstrator.

	Application	Network	Humans / Machinery
Application	Monitoring end-to-end dependability (4.5) DTs for Emergent Intelligence (5.7)	Compute and Artificial Intelligence (AI) resource allocation (3.3, 3.4) Error identification (4.2)	Communication-Control-Codesign (4.1)
Network		Interference management, link scheduling (3.1, 4.4) Ambient backscatter communications (3.5) RAN function distribution (3.6) Distributed Massive MIMO (4.3) Control/data plane guarantees (4.5)	Trajectory planning (3.2, 4.4) Human presence (5.3) Network-aware digital twin (4.4, 5.6)
Humans / Machinery			Novel HMIs and Human in the Loop (5.2, 5.4) Collaborative robots (5.5)

2.2 Updates on use cases and targeted KPIs/KVIs

Hexa-X recently published an update of use cases, their grouping into use case families and the targeted KPIs/KVIs based on the input gathered from technical work packages in [Hexa-X D1.3]. In the following, we briefly discuss updates of KPIs and KVIs that are relevant to our work package. We provide a first overview of how our contributions address these KPIs and KVIs based on the initial solutions discussed in this deliverable. A full assessment of solutions against target KPIs for individual use cases is planned for our final deliverable D7.3 due in May 2023.

[Hexa-X D1.3] groups KPIs by services and capabilities offered by the 6G system: communication, AI/computation, and localization/sensing. We use this structure in Table 3 to map our contributions against KPIs. For the definitions of the respective KPIs, please refer to [Hexa-X D1.3], Section 2.2.1.

Table 3: Contributions mapped to targeted KPIs.

		KPI	Contributions
Communication	Dependability	Availability	Understand and model the impact of packet losses etc. on application performance with Communication(-Computation)-Control-Codesign (4.1), with the goal to achieve high application productivity Quantification of end-to-end dependability (4.5) *All contributions mentioned under "Quality of Service (QoS) Attributes"*
		Reliability	
		Safety	No dedicated contributions
		Integrity	No dedicated contributions
		Maintainability	Observability with dependability monitoring (4.5) Mechanisms for error identification (4.2)
	QoS Attributes	Service latency	Deterministic latency with control and data plane guarantees (4.5)
		Data rate	Increasing the efficiency of Radio Resource Management (RRM) with Digital Twin (4.4) Dependability in Massive MIMO (4.3) Utilizing ambient backscatter communications for resource constrained devices (3.5) Interference management (3.1)
		Resource constraints	
		Scalability	
AI and computation	Dependability Attributes	Agent availability	Inclusion of "Computation" in Communication-Control-Codesign (CoCoCo) (4.1) Resource provisioning for Federated Learning in IoT (3.4)
		Agent reliability	
		Safety	No dedicated contributions
		Integrity	No dedicated contributions
		Maintainability	Observability with dependability monitoring (4.5) Mechanisms for error identification (4.2)
	QoS Attributes	AI service RTT	Optimal resource allocation and redistribution (3.3) Resource provisioning for Federated Learning in IoT (3.4)

			Communication-Computation-Control-Codesign (4.1)
		Inferencing accuracy	No dedicated contributions
		Interpretability level	No dedicated contributions
		Training/model transfer latency	Optimal resource allocation and redistribution (3.3) Resource provisioning for Federated Learning in IoT (3.4) Digital Twins for Emergent Intelligence (5.7)
		Resource constraints	
		Scalability	
Localisation and Sensing		Utilization of location information in digital twins (e.g., for trajectory optimization) to allow for more efficient resource utilization (3.2, 4.4). Modelling the impact of human presence, potentially augmented by sensing capabilities of 6G (5.3). Selected aspects are studied in a collaboration between WP3 and WP7 on *location and sensing enhanced services* as described in [Hexa-X D3.1], Section 5.	

With respect to KVIs, the categorization of use cases presented in [Hexa-X D7.1] into *dependability for I4.0* and *sustainable coverage in IoT* remains valid, addressing mainly the key values *trustworthiness* and *sustainability*. Table 4 provides an initial mapping of technical contributions discussed in this deliverable to KVI areas, without going into details on the underlying use cases. Use case contributions to KVI areas are being discussed in [Hexa-X D1.3], Section 2.3.

Table 4: Mapping of contributions to Key Value Indicators (KVIs).

KVI area	Contributions	Remarks
Sustainability	Ambient backscatter communication (3.5)	Novel zero-energy devices for massive IoT scenarios (e.g., earth monitor)
	Efficient resource allocation (Sec. 3)	Efficient utilization of infrastructure, adapted to current load and conditions (c.f. flexibility KVI)
Trustworthiness	Dependability-related contributions (Sec. 4)	Increased and observable/quantifiable dependability is expected to contribute to the overall level of trust as an indicator of trustworthiness [Hexa-X D1.3]
	Trustworthy Digital Twin platform (5.1, 5.6, 5.7)	Privacy-preserving collaboration among digital twins, benefiting from novel 6G capabilities (e.g., localization, sensing)
Inclusiveness	Novel HMIs (5.2) and interaction with Digital Twins (5.4, 5.5)	Enable remote interaction, enable inclusion of a more diverse (remote/on-site) workforce. Reduced human exposure to hazardous/dangerous situations.
Flexibility	Flexible resource allocation (Sec. 3)	Mechanisms to adapt to changing requirements, mobility, device constraints, …

Special-purpose functionality and the resulting solutions rely on technical enablers from other work packages and are part of the overall end-to-end architecture presented in [Hexa-X D1.3]. When discussing performance targets, achievable results therefore depend on those enablers and architectural principles as well. In the following, we discuss relevant relations to architectural principles and other technical work packages that will have an impact on the overall performance. We will report updates on these relations in D7.3, available in May 2023.

2.3 Relation to end-to-end architecture, architectural principles and technical enablers

Contributions for special-purpose functionality need to be considered in the context of the overall end-to-end architecture and its enablers. We provide an initial mapping of our proposed functionality to the 6G architectural principles and enablers, thereby illustrating the role of special-purpose functionality in future 6G systems. *Architectural principles* for 6G, as summarized in [Hexa-X D1.3], Section 3.1.1 and further detailed in [Hexa-X D5.1] are:

Exposure of new and existing capabilities – especially relevant in the digital twin ecosystem discussed in this deliverable in Section 5 and when it comes to the utilization of localization and sensing information in the digital twin for, e.g., trajectory optimization (Sec. 3.2).

Designed for (closed loop) automation – enables the execution of resource allocation algorithms and methods as discussed in Section 3, for efficient placement of network and application elements. Not only on a global scale, but also for local networks or networks-in-networks and compute/AI resources on resource-constrained devices and networks.

Extensibility and flexibility to different topologies – supporting the full range: from models for In-X networks (or network-of-networks) in factory environments that can be utilized for interference management and link scheduling (Sec. 3.1), to support for ambient backscatter communications on resource constrained zero-energy devices (Sec. 3.5).

Scalability – a consequence of flexible resource assignment and efficient use of available resources, for example as a consequence of careful Communication-Computation-Control-Codesign (Sec. 4.1) or by benefitting from more detailed digital twins (Sec. 4.4).

Resilience and availability – a focus of the work presented in this deliverable, especially in Section 4. Resilience and availability is increased by understanding the impact of errors on application productivity, and by extending dependability to an end-to-end perspective (Sec. 4.2 and 4.5). As previously discussed for scalability, an increased resilience is also a consequence of flexible resource assignment and the efficient use of resources (Sec. 4.1, 4.4).

Exposed interfaces are service-based – this is expected to further ease the integration of network services and capabilities into the digital twin ecosystem, thereby allowing digital twins to benefit from additional network awareness and insights (Sec. 3.2, 4.4, 5.6). Additionally, this principle should allow easier extension of networks with specialized services offering some of the functionality discussed in this deliverable (e.g., on resource allocation, Sec. 3).

Separation of concerns of network functions – expected to further allow tailored fitting of deployed functions to the respective scenario, especially in local networks (or networks of networks). However, this is not in the focus of the work reported in this deliverable.

Network simplification in comparison to previous generations – considered especially relevant for the management and operation of local networks or networks-in-networks. Complexity also affects the ability of the digital twin to capture all relevant network-related aspects and the ability of humans to interact with the network and its digital twin through novel HMIs (Sec. 5).

With respect to *technical enablers* for the end-to-end architecture as discussed in [Hexa-X D1.3], Section 3, the functionality and solutions proposed in this deliverable benefit from advances and further alignment in the areas shown in Table 5. The table summarizes all areas and enablers and provides pointers to the respective deliverables. As work regarding these enablers is still ongoing in parallel to activity in our work package, there is need for continued alignment. Respective activities for relevant enablers are also summarized in the table. Progress on these alignments will again be reported towards the end of the project in deliverable D7.3, due in May 2023.

Table 5: Related technical enablers.

Area	Enabler	Details	Activity
RAN	High data rate radio links	[Hexa-X D2.1], [Hexa-X D2.2]	Observe performance bounds and constraints
	Distributed large MIMO	[Hexa-X D2.2]	Contributing with dependability in distributed massive MIMO (Sec. 4.3)
	Localization and sensing	[Hexa-X D3.1]	Formulate requirements and observe performance bounds and constraints, especially regarding utilization in the digital twin (Sec. 4.4., 5.3, 5.6)
Intelligent network	UE and network programmability	[Hexa-X D5.1]	UEs not yet explicitly considered - can be an enabler for novel HMI functionality or retrofitting.
	Network automation	[Hexa-X D5.1]	C.f. continuum management and orchestration.
	AI and AI as a Service	[Hexa-X D5.1], [Hexa-X D4.1]	Special-purpose case of federated learning in IoT (Sec. 3.4), AI and Emergent Intelligence in digital twins (Sec. 5.7)
	Dynamic function placement	[Hexa-X D5.1]	Utilized as enabler for resource allocation in challenging environments (Sec. 3), impact on dependability (e.g., Sec. 4.5)
Flexible network	Mobility solutions	[Hexa-X D5.1]	HetNet approach, considered for In-X networks in factory environments (Sec. 3.1) and works on D-MIMO (Sec. 4.3)
	Campus network	[Hexa-X D5.1]	Capability exposure and mutual benefit with collaborating digital twins (Sec. 5.6), utilization of CoCoCo (Sec. 4.1), resource allocation strategies for compute (Sec. 3)
	Mesh / Device-to-Device	[Hexa-X D5.1]	In-X networks in factories (Sec. 3.1)
	Edge cloud integration	[Hexa-X D5.1]	Enabler for most resource allocation strategies (Sec. 3) and local placement of functions in low-latency and high-dependability scenarios
Efficient network	Cloud and service-based architecture	[Hexa-X D5.1]	Reduced dependencies between network functions enables more flexible placement and, thereby, adaptation to latency requirements. Relevant for industrial scenarios with increased dependability requirements (Sec. 5.6)
	Compute-as-a-Service	[Hexa-X D5.1]	Availability of (trustworthy) compute capabilities for the execution of digital twins (Sec. 5). Can be enriched with allocation strategies, e.g., for federated learning in IoT scenarios (Sec. 3.4)
Service management	Continuum management and orchestration	[Hexa-X D6.1]	Formulation of an ecosystem of digital twins to allow cross-domain optimization and collaboration in a privacy-preserving fashion (Sec. 5). Enabler for resource allocation, e.g., in In-X networks and networks-of-networks (Sec. 3.1)
	AI-driven orchestration	[Hexa-X D6.1]	Exposure of local (domain-)knowledge through network-aware collaborating digital twins to aid in overall resource coordination and management (Sec. 4.4, 5.6)

2.4 Security considerations

Upon the high security requirement in I4.0 scenario, many use cases discussed in this document may have critical concerns about security and privacy, as summarized in Table 6. The identified security concerns shall be further analysed and addressed in alignment with the Hexa-X work on security and trustworthiness presented in [Hexa-X-D1.3]. Progress on these alignments, including the proposal of solutions regarding the identified security concerns. Progress on these alignments will again be reported towards the end of the project in deliverable D7.3, due in May 2023.

Table 6: Security concerns in the discussed use cases.

Use Case	Concerns
Novel HMI for mobile human-machine and human-cyber-physical-environment (CPE) interaction (Sec 5.2)	1. Privacy-critical biometric information is transmitted over the radio link, exposed to the risk of wiretapping 2. Privacy-critical biometric information may be unauthorized exploited 3. Multi-sensory feedback from machine to human may be manipulated by unauthorized to generate malicious (and even dangerous) impacts to human users
Digital twins for Emergent Intelligence (5.7)	1. Data exchange between different agents may be wiretapped 2. Malicious users may pretend as normal agents to inject fake information into the Emergent Intelligence (EI) system
Usage of shared infrastructure (Section 4)	Public infrastructure may not be trusted because data might be gathered and leaked to third parties. This includes both providers of communication services as well as providers of computation services. Currently, the trend is to deploy and manage own infrastructure.

3 Flexible resource allocation in challenging environments

This section reports the advancements in the use cases investigated in WP7 and dealing with the need for ultra-flexible resource allocation caused by the bandwidth crunch in traditional radio bands.

The use cases and procedures presented focus on challenging environments populated by mobile devices with special requirements and in need of coverage. One example, in Section 3.1, is the flexible resource allocation for cellular-system-based communications in factory environment, involving generic machines including vehicles. Generally, this use case looks at the coexistence of three distinct types of communications with heterogeneous coverage areas. The challenges of the factory environment are further investigated in Section 3.2, leveraging location information and anticipated trajectories of moving, connected machines like Automated Guide Vehicles (AGVs) to flexibly allocate resources and to update the radio map with static and moving obstacles on the factory floor. Section 3.3 takes instead a different look at the challenges in this environment by designing algorithms for functionality placement and redistribution of tasks in case of malfunctions of robotic units and edge nodes. Moving outside the factory, Section 3.4 tackles the proliferation of devices, especially IoT devices like sensors, cameras and actuators in urban areas at large, analysing the impact on network resources of distributed AI techniques that require frequent interactions between AI agents. The crowding of the spectrum in urban areas can also be addressed by recycling radio waves, in what is called "Ambient backscatter communication", a technique presented in Section 3.5. Finally, a comprehensive solution to maximise the available resources in dynamic, challenging environments is presented in Section 3.6: it entails a flexible, adaptable and scalable RAN functional split, which can be implemented in hardware and software.

3.1 In^2-X communication in factory environments

One significant use case of 6G flexible resource allocation is the cellular-system-based In^2-X communications in factory environment, where In^2 refers to both the prefixes "intra-" and "inter-", and the X stands for generic machines including vehicles. Generally, this use case holds the factory owner accountable for communications involving the machines, as illustrated in Figure 3-1. More specifically, it encompasses the following three types of communication depicted in Figure 3-2:

1. The intra-X communication, i.e., the data exchange between different components contained within the same machine. These communications must be cost effective while maintaining a high level of reliability, with a small coverage area limited within the machine. The traffic pattern of intra-X is generally deterministic and can be predicted by the manufacturer. The communication devices must be integrated into the machine by manufacturer and operate without the use of a subscriber identification module (SIM)card.
2. The inter-X communication between different machines. These communications typically have an intermediate range of coverage. A high degree of flexibility is required, as the characteristics of the radio channels may frequently vary upon redesign of the machines and the production lines. The manufacturer is commonly unaware of the pattern of inter-X traffic, but it is generally polycyclic, determined by multiple factors, such as the selection of machines, the manufacturing process flow, among others. Therefore, AI-driven solutions are required to autonomously integrate new APs in response to link requirements and to monitor traffic against potential anomalies.
3. The X-to-infrastructure (X2I) communication between the machines and the factory infrastructure. Generally, coverage of the entire factory is required. It shall be noted that the infrastructure consists of numerous moving facilities such as AGVs, and even service staff can be considered as part of the factory infrastructure. These facilities and human participators generate complex patterns of mobility, which also elevates the complexity of the X2I traffic pattern beyond polycyclic. Not only AI-driven approaches are required to address the challenges of traffic pattern recognition and prediction, but also, to conduct radio analysis for predictive maintenance, as well as to detect intrusion and network anomalies.

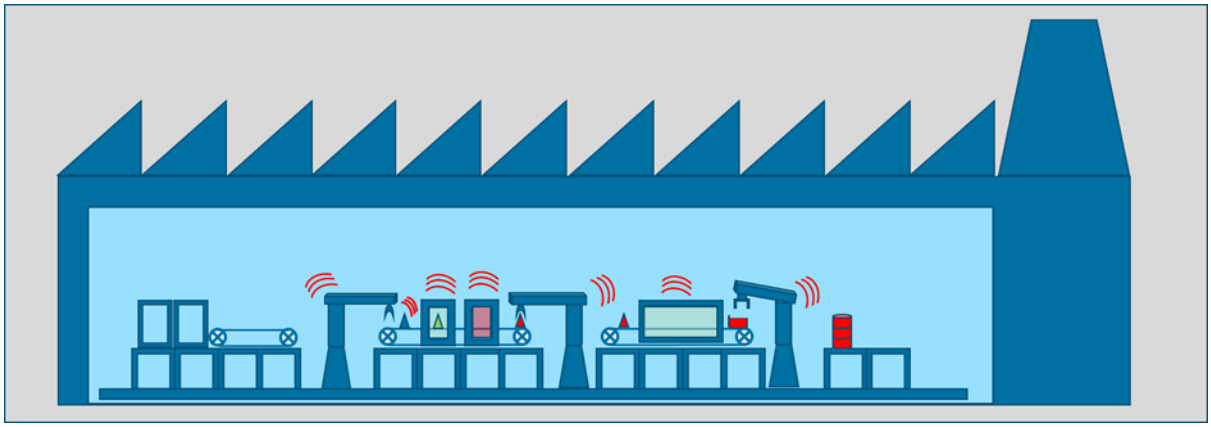


Figure 3-1: The In^2-X communication scenario in factory environment.

The coexistence of three distinct types of In^2-X communications with heterogeneous coverage areas and traffic patterns challenges conventional approaches of interference management and link scheduling. It is seeking support from multiple technical enablers, including AI-driven automated self-integration of new APs, underlay networks, efficient radio-resource management with radio-aware DTs (c.f. Section 4.4), and integration in machines. To pave the way towards such AI solutions, we build the topology model of the In^2-X use case, as illustrated in Figure 3-2. The topology is characterized by a cascaded network in network structure (in the example in Figure 3-2: outside network, in factory hall network, in machine network). The three stages of networks are successively overlayed on top of the other, with each successor as the "outer network" of its predecessor (the "inner network").

The inner network, regardless of its own topology (i.e. either with a central node or based on a device-to-device setup), is integrated into the next outer network and is, in general, overall power controlled by the next outer network in order to avoid an unacceptable degradation of the next outer network by interference from the inner network. However, it shall also be capable of autonomously working upon necessity, e.g. when the X2I communication fails, as suggested in [BBA+21]

In addition, the inner network should for safety and resilience reasons be able to operate independently, might be isolated from the next outer network by an extra isolation (e.g., caused by a vehicle cabin), and in many scenarios is mobile. At least one node of the inner network is connected to the next outer network.

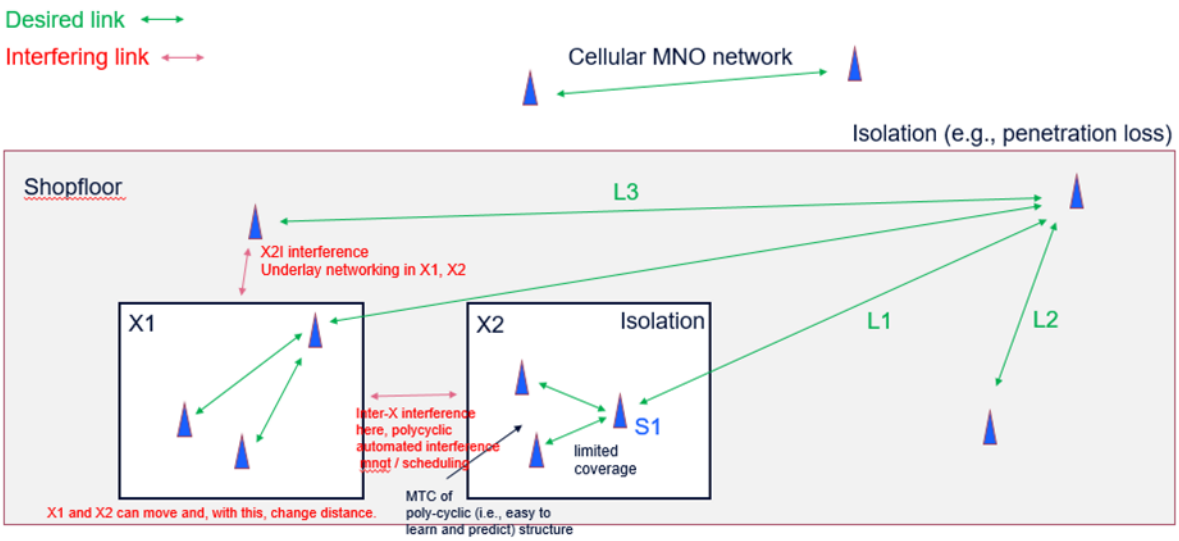


Figure 3-2: The In^2-X communication topology model.

3.2 Radio-aware Trajectory planning

3.2.1 Motivation

6G industrial use cases like "Digital Twin for Manufacturing" can be extremely demanding in terms of required latency and reliability for mission-critical services (like industrial ethernet over the air). At the same time, the use cases can require other services, such as extremely high-quality video, augmented reality (AR) / virtual reality (VR), which need very high bitrates and bandwidths. These diverse demands also vary a lot based on the time and location, depending on the location of stationary and mobile obstacles. On the other hand, 6G will offer high amount of new spectrum (for example sub-THz) on top of the old spectrum. To handle all this, resources should be allocated very flexibly in temporal, frequency and spatial domains.

The service demands and their KPIs are often very local. For example, one production cell in a factory can have very high connectivity demands while the neighbouring production cell does not. Furthermore, the User Equipments (UEs) are often moving: in a factory environment there are often moving machines like AGVs requiring connectivity. Other examples include slowly moving production cells in automotive manufacturing industry, drones and Unmanned Aerial Vehicles (UAVs) in various use cases, or wireless cargo handling vehicles in port automation [UVK+21].

While the classical KPIs are focused on radio performance, in 6G we have also new KPIs/KVIs which have much broader targets for the whole society, for example sustainability and energy efficiency [Hexa-X D1.3]. In 4G and 5G, the latter was mostly just about UE energy consumption to save battery life and reducing always-on broadcasted information in favour of on-demand transmission. In 6G, the energy efficiency is seen as a target itself, the whole end-to-end network energy consumption has to be reduced and even performance can be sacrificed in order to save energy.

In this section we show how we can utilize UE location information and anticipated trajectories to flexibly allocate resources in challenging industrial environments. One key enabler for this is our use case where the UEs are not associated with humans but with machines. This enables much greater (even full) predictability of user location and mobility, and we can use mobility to our advantage. However, the same – typically AI enabled – mechanisms used for predicting the UE movement will also be able to identify not yet known movement pattern or anomalies in movements that can indicate accidents or even malicious manipulation of systems.

3.2.2 Radio awareness and trajectory planning

In section 4.4.2 we discuss about radio-aware DT where we do not just model a virtual copy of the network but also make it radio-aware so that the DT knows all the transmitting/receiving radio nodes in the network and is also able to predict radio conditions in certain location at certain time (also in the near future). When we are in a controlled environment, we can know exactly where the UEs are and also in many cases where they will be in the near future. As an example, AGVs moving on factory floor are typically using fixed trajectories and their routes are planned in advance (at least within some time window in the near future). Thus, we can know where all the UEs are at certain time in the near future (e.g., after 5 seconds). This enables us to do many things pro-actively and in an anticipatory manner, so for example in case of Ultra Reliable Low Latency Communications (URLLC), we do not wait for errors to happen (even 1 or two errors can be critical, and we might not have delay budget for retransmissions) but we can for example allocate more or better radio resources (time/frequency channels, beams, different frequency band etc.) in advance *before* errors happen.

Another new degree of freedom is that we can also control the mobility. For example, we can make a moving AGV or drone take a different route with the same destination. Even if the new route were longer, we would be able to satisfy some KPIs better when using the new route. For example, we can reach the reliability and latency target since the new route has much better Signal-to-Interference-plus-Noise Ratio (SINR) compared to the shortest default path. Or we can have much shorter file transfer rates with the new route since we can connect to a very high-capacity sub-THz access point along our new route. Or we can save energy by being able to reduce the transmission power of involved devices

(of course, the re-routing could then mean that more energy is consumed for mobility). We can also trade off performance with energy efficiency by choosing our optimization criteria as we show in next section.

In general, we propose to use a Radio-aware Digital Twin, which can be seen also as one example of CoCoCo (Section 4.1). This DT uses a Radio Environment Map which is not static but is constantly kept up-to-date and it has also a prediction engine which can predict radio conditions in certain location over some time horizon. With controlled machine-type mobility (UEs are associated with AGVs or drones/UAVs) we can plan the routes of the UEs based on the desired optimization criteria, for example total aggregated throughput, latency, reliability etc. We can also predict the interaction of the moving UE with other users in the network in terms of generated interference, radio resource consumption etc. The whole network operation can be jointly optimized based on selected criteria and selected boundary conditions.

3.2.3 Exemplary results with radio-aware trajectory planning with UAV

As an example of radio-aware trajectory planning, we show some system simulation results with UAV which has a mission (for example for video surveillance) to travel from Point A to Point B. The UAV is moving in a macro cell network and has some constraints for the mission (boundary conditions):

- The UAV has an energy constraint due to limited battery capacity
- The UAV has a time constraint, the target must be reached within certain time budget
- The UAV has a minimum throughput requirement

Within these boundary conditions, the target is to for example maximize the UE total aggregated throughput or minimize energy consumption along its route. The solution assumes Radio Aware Digital Twin with radio environment map (REM) prediction based on simple pathloss estimates excluding interference. The energy consumption of the drone is modelled accurately and includes the energy required for hovering, turning, and flying at a certain speed. This model also factors in the wind direction relative to the UAV's position.

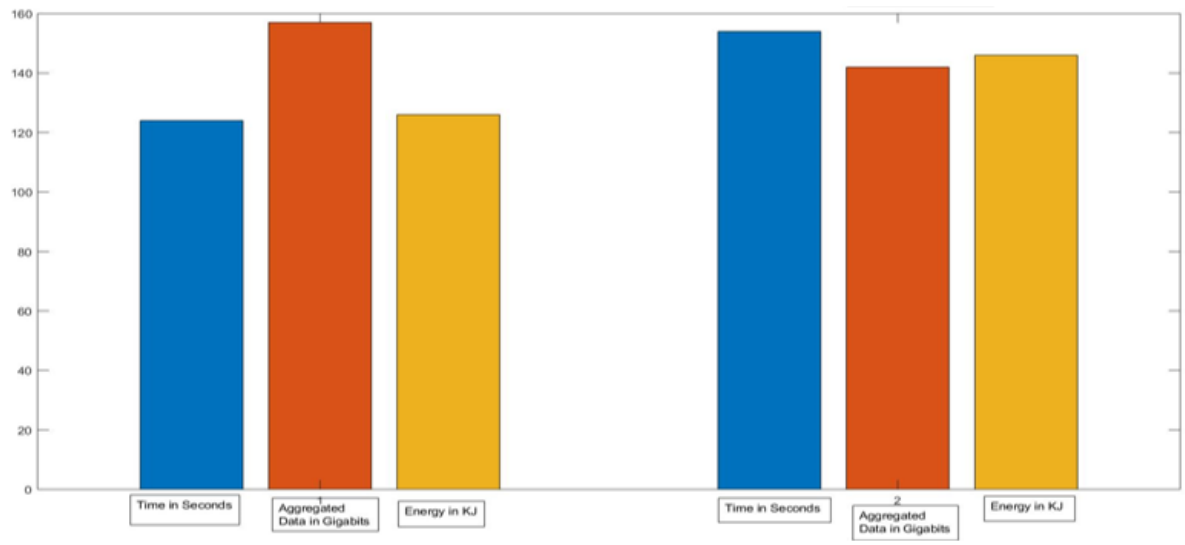


Figure 3-3: Comparison between the optimal path that maximizes the weighted sum of the data and energy consumption (left) and the solution from a greedy algorithm (right).

The figure above shows results with two different algorithms. On the left: the optimal solution that uses dynamic programming to optimize a weighted sum of the aggregated data and the energy consumption of the UAV. On the right: a simple greedy algorithm where the direction and speed of travel is based only on the currently observed path loss (as a proxy for data rate) and energy consumption. Compared

to the baseline (direct path), the optimal algorithm provided 65% more aggregated data download at the cost of 25% increase flight time and energy consumption. Note that in this case, the energy consumption is dominated by the power needed to fly the UAV (for example – the power consumption for hovering is of the order of 100s of Watts, whereas the power consumption of a radio transmitter is less than 10 Watt), thus a straight path is always the best from the energy consumption point of view.

3.2.4 Flexible radio mapping

During operation in a dynamic environment, the radio map utilized for radio-aware trajectory planning requires updates in case of physical reconfigurations of the space. One can distinguish between two types of reconfigurations: planned and unplanned. Planned reconfigurations include changing the layout of a modular production cell in an I4.0 scenario or deploying new network infrastructure elements such as additional radio units, but also altering predetermined trajectories of an automated guided vehicle. Unplanned reconfigurations are, for example, interference or blockage introduced by goods being moved around or introduced by humans moving on the factory floor. The effects of planned reconfigurations can (at least to some extent) be modelled and included in the re-calculation of a radio map during operation. However, the effects of unplanned reconfigurations cannot fully be considered beforehand (one approach that utilizes knowledge about humans moving on a factory floor to update a radio map is discussed in Sec. 5.3) and, therefore, need to be detected in the deployment.

To this end, we tag measurements of network and application metrics with their respective geo-location and collect the data in a time/geo series database (details on the general setup are provided in Section 4.5). Measurements are taken by both static and mobile sensors. For static sensors, time series analysis can be utilized to detect unexpected changes and trigger (manual) inspection. For mobile sensors, data along trajectories is captured and can be compared with similar trajectories – in a factory setting, having multiple measurements of the same trajectory is a reasonable assumption. In addition to the time series analysis, measured data on radio link quality can be compared to the pre-calculated radio map of the environment to detect clusters of outliers and trigger human intervention or automated adaptation. This data can be utilized to continuously improve existing radio map and mobility models.

3.3 Optimal resource allocation and redistribution of functionalities in industrial environments

Motivation

In this section, we describe the preliminary work on Functionality Allocation algorithm design for functionality placement and redistribution for ensuring efficient operation of a robotic system. We assume a system with various robotic units conducting various tasks and cooperating with each other in industrial environments.

In the case of an unexpected situation in industrial environments, like the case where a task / functionality is in pending mode or is executed slowly, it is crucial to reallocate this malfunctioning task to one of the other nodes / robots and avoid a possible downtime. Accordingly, in the case where a robotic device goes out of order, there should be a component / algorithm responsible for redistributing the functionalities to the rest of the nodes/robots/units in an orchestrated manner with minimum cost. Hence, we are studying the following functionality allocation problem to ensure reliable industrial automations.

Problem Statement

We assume a set of Functional Entities (FEs) e.g., tasks, jobs, services, and a set of Hosting Entities (HEs) e.g., robotic units, edge nodes, core nodes. Each FE has a computational load and each HE has some capabilities. These capabilities are the maximum computational load (related to the number of available CPU cores, RAM and disk storage), the battery level (if applicable), and the functionality types that can be supported by each HE. By this last capability we mean the set of FEs that each HE can support (e.g., a robotic unit with camera can support live streaming but cannot support grasping an

object if robotic arm is not available). Moreover, a functional graph is considered with nodes corresponding to the FEs and edges connecting interacting FEs weighted according to the amount of data transferred between FEs. Finally, a system layout graph is considered consisting of nodes corresponding to the available HEs and the communicational channels among them with different communication costs and maximum capacities.

Our objective is the allocation of FEs to HEs. Specifically, we are looking for the minimum cost allocation that satisfies a set of performance constraints.

The objective function is associated with:

- the cost of utilizing a HE (first term of the equation), mostly related to the battery level of the HE, if applicable. This cost takes higher values when battery is low and close to zero values when it is fully charged or when the HE is not battery-powered.
- the power consumption cost of running a FE on a HE (second term of the equation).
- the computational cost of running a FE on a HE (third term of the equation), which is related to the computational load of the FE and the maximum computational load of the HE,
- the cost (latency) imposed by the communication among HEs (last term of the equation), which may be related to the availability of the HE to conduct the job (i.e., standby mode etc.).

The constraints of our problem address the following aspects:

- all FEs need to be assigned to the available HEs.
- all capabilities of HEs need to be respected.
- the capacity constraint of each communicational link should be respected.

Problem Formulation

The notations used for functionality allocation problem formulation can be found in Table 7.

Table 7: Functionality allocation notations.

Notation	Definition	Notation	Definition
n, m	Total number of FEs and HEs respectively	$k_{i,i'}$	Weights of interacting FEs F_i and $F_{i'}$
$i,\ i'$	Indexes of FEs	$G_H = (H, L)$	System layout graph with nodes HEs and computational channels L
$j,\ j'$	Indexes of HEs	$r_{j,j'}$	Cost of communication between HEs H_j and $H_{j'}$
$F = \{F_1, \ldots, F_i, \ldots, F_n\}$	Set of FEs	$cap_{j,j'}$	Maximum capacity of the links among HEs H_j and $H_{j'}$
$H = \{H_1, \ldots, H_j, \ldots, H_m\}$	Set of HEs.	$e_{i,j}$	Power consumption cost of running FE F_i on HE H_j
φ_i	Computational load of FE F_i	$w_{i,j}$	Binary constant showing if FE F_i can be assigned to HE H_j in terms of functionality types that can be supported by the HE
c_j	Maximum computational load of HE H_j	A={ A_1 , …, A_j, …, A_m}	Collection of sets of FEs that will be assigned to the HEs

b_j	Cost of utilizing HE H_j (related to the battery level)	y_j	Decision variable that takes 1(0) depending on whether HE H_j is (is not) utilized
$s_{i,j}$	Computational cost of running FE F_i on HE H_j	$x_{i,j}$	Decision variable that takes 1(0) depending on whether FE F_i is (is not) assigned to HE H_j
$G_f = (F, K)$	Functional graph with nodes FEs and edges K	$z_{j,j'}$	Decision variable that takes 1(0) depending on whether HE H_j and $H_{j'}$ are (are not) communicating

We introduce the constants $w_{i,j}$, to indicate the feasibility of assigning a FE to a HE in terms of functionality types that can be supported by the HE (takes 1 when HE can support the FE F_i, and 0 otherwise). We assume that all FEs can be supported by at least one HE ($\sum_{j=1}^{m} w_{i,j} \geq 1,\ \forall\, 1 \leq i \leq n,\ i \in \mathbb{N}$).

Moreover, we introduce the set of decision variables y_j, that take the value 1 or 0 depending on whether the HE H_j is utilized or not. We also define the set of decision variables $x_{i,j}$ to describe the allocation of FEs to HEs, taking the values 1 or 0 depending on whether FE F_i is assigned to HE H_j or not. Finally, we define the set of decision variables $z_{j,j'}$ to describe the communication among HEs, taking the values 1 or 0 depending on whether H_j and $H_{j'}$ are communicating due to communicating FEs F_i and $F_{i'}$ assigned to them respectively or not.

The problem of obtaining the set A of FEs F that will be assigned to each HE H_j may be reduced to the following problem:

$$\min_{y,x,z} \quad \sum_{j=1}^{m} y_j\, b_j + \sum_{i=1}^{n}\sum_{j=1}^{m}\left[x_{i,j}\ \cdot e_{i,j}\right] + \sum_{i=1}^{n}\sum_{j=1}^{m}\left[x_{i,j}\ \cdot s_{i,j}\right] + \sum_{j=1}^{m}\sum_{j'=1,\, j'\neq j}^{m}\left[z_{j,j'} \cdot r_{j,j'}\right]$$

Subject to:

- $\sum_{j=1}^{m} x_{i,j} = 1,\ \forall\, 1 \leq i \leq n,\ i \in \mathbb{N}$, all FEs are allocated,
- $\sum_{i=1}^{n}\left[x_{i,j} \cdot \varphi_i\right] \leq c_j,\ \forall\, 1 \leq j \leq m,\ j \in \mathbb{N}$, the maximum computational load of the HEs is respected.
- $y_j \geq x_{i,j},\ \forall\, 1 \leq i \leq n,\ 1 \leq j \leq m,\ i,j \in \mathbb{N}$, all HEs that are utilized ($y_j = 1$) have at least one FE assigned on them.
- $x_{i,j} \leq w_{i,j},\ \forall\, 1 \leq i \leq n,\ 1 \leq j \leq m,\ i,j \in \mathbb{N}$, the feasibility of assigning a FE to a HE in terms of functionality types that can be supported by the HE is respected.
- $\sum_{i=1}^{n}\sum_{i'=1}^{n} z_{j,j'} \cdot k_{i,i'} \leq cap_{j,j'}\ \forall\, 1 \leq j,j' \leq m,\ j,j' \in \mathbb{N}$ where $i \neq i'$ and $j \neq j'$, the maximum capacity of the links among HEs is respected.
- $z_{j,j'} \geq x_{i,j} + x_{i',j'} - 1$, $\forall\, 1 \leq i,i' \leq n,\ 1 \leq j,j' \leq m$, $i,i',j,j' \in \mathbb{N}$ where $i \neq i',\ j \neq j'\ and\ k_{i,i'} \neq 0$, the communicating HEs should have communicating FEs assigned on them.

The described functionality allocation problem is a mixed integer programming problem. It will be modelled and solved with the use of the open-source Python-MIP[2] library, it will be containerized using

[2] Python-MIP, "Mip optimizer for python," 2019. Available from: https://python-mip.com

Docker[3] and it will be placed in the central infrastructure. Also, an application programming interface (API) documentation will be created for this algorithm by using the Swagger[4] open-source tool. The connection with the overall architecture is described in Section 4.2.

3.4 Resource provisioning for Federated Learning in IoT

The proliferation of IoT devices in urban areas, which already number in the thousands and are expected to grow dramatically in coming years, brings about potential network issues during communication between the devices and the server collecting IoT data by introducing delay, packet loss and bandwidth congestion. In view of the use of distributed, collaborative approaches to gain insights into data generated by IoT devices, it is expected that AI techniques, and especially Federated Learning techniques, will become commonplace. It is thus important to analyse whether the network issues cited above may be exacerbated by a Federated Learning (FL) approach; especially since it is a synchronous process so also a little delay in a client multiplied by each round of the algorithm can cause a very high delay in the whole process.

We decided to use two different frameworks in order to be able to compare the results. In particular the first one is FLOWER [BTM+20], i.e., an open-source framework, platform independent, easy to understand and customize. The other one is called IBMFL [LBT+20]: a black-box Python framework realized by IBM where only the dataset and the convolutional neural network (CNN) model are customizable.

3.4.1 Scenario

For all the experiments presented in Section 3.4.2 the configuration used consists of one server and some clients, all running on different Linux Docker containers on the same virtual machine. Before each experiment, the Transmission Control Protocol (TCP) segmentation offload is disabled and the bandwidth limit is set to 1 Gbit/s. The bandwidth limit and other delay or packet loss are introduced using NetEm with traffic control on Linux. For the dataset we decided to use MNIST [Den12], a simple dataset containing images of handwritten numbers from 0 to 9, but we have replicated all the results also for a bigger dataset as CIFAR-10 [CIFAR], that contains images belonging to 10 different classes.

3.4.2 Experimental Results

We now describe the main outcome obtained with the two frameworks described above. In particular, first we study the federated learning approach in an ideal condition, and, next we investigate how network issues influence the aforementioned approach, showing a simple run of the Federated Learning algorithm between 2 client and a server, without any delay on the network and using the same CNN over the same dataset.

The two graphs in Figures 3-4 and 3-5 show the traffic captured on the server, in particular the green line refers to the packets sent by the server to the two clients, while the blue and red ones refer to the packets received from the clients. As the test is done in the best possible conditions, i.e., large bandwidth, no delay, same amount of available resources on the two clients, the two curves of the clients are completely overlapped, a behaviour observed both on the FLOWER framework and on the IBMFL framework. It must be remarked that in the figures it may appear counterintuitive that the peaks of sent packets occur right after the peaks of received packets. However, it depends on the operations of the two frameworks. In FLOWER one can see a first green peak when the server sends to the clients the initial global parameters, then it receives the local parameters from the clients and re-send the updated global parameters. Instead, in IBMFL, one cannot see the first peak because the framework does not

[3] Docker. What is a Container? Available from: https://www.docker.com/resources/what-container.

[4] Swagger. Swagger: API Development for Everyone. Available from: www.swagger.io

send the initial global parameters but it just sends a packet asking the clients to start their data fitting process.

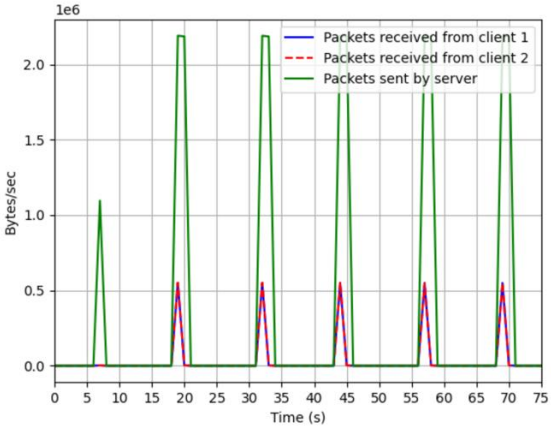


Figure 3-4: FLOWER framework. 2 clients and 1 server, no delay and bandwidth limit 1 Gbit/s

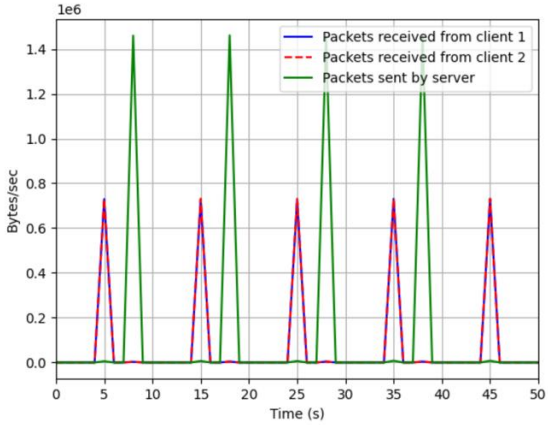


Figure 3-5: IBMFL framework. 2 clients and 1 server, no delay and bandwidth limit 1 Gbit/s

In the next step, we add a delay to one or more clients. The first test (Figures 3-6 and 3-7) consists in the replication of the ideal test but this time introducing a bi-directional delay of 150 ms to only one of the two clients. They show the running of the Federated Learning algorithm between 2 clients and a server; client 2, red line in the graphs, has a bi-directional delay of 150 ms, and both clients are using the same CNN over the same dataset. Like before the graphs represent the traffic captured on the server but unlike before the red and blue lines are not overlapping anymore, due to the delay of the client in red. Also, the performance in terms of total execution time is quite different, for FLOWER due to the delay, the total execution time has gone from 63 s of the ideal test to 74 s, while for IBMFL from 50 to 55 seconds.

Finally, Figure 3-8 shows a different configuration of the federated learning approach: in this test there are 4 clients and a server, three of them have a bi-directional delay of 50 ms while the fourth one has 150 ms of delay.

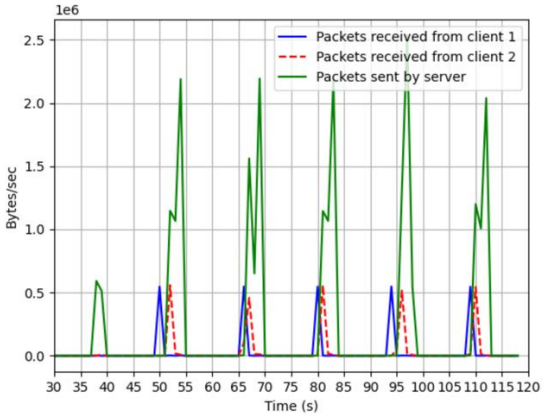


Figure 3-6: FLOWER framework. 2 clients and 1 server, client 2 (red) with 150 ms bidirectional delay and bandwidth limit 1 Gbit/s

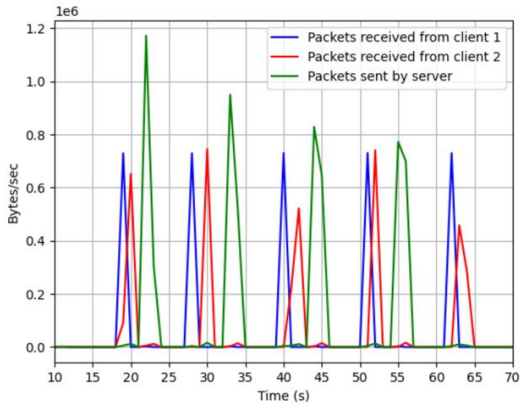


Figure 3-7: IBMFL framework. 2 clients and 1 server, client 2 (red) with 150 ms bidirectional delay and bandwidth limit 1 Gbit/s

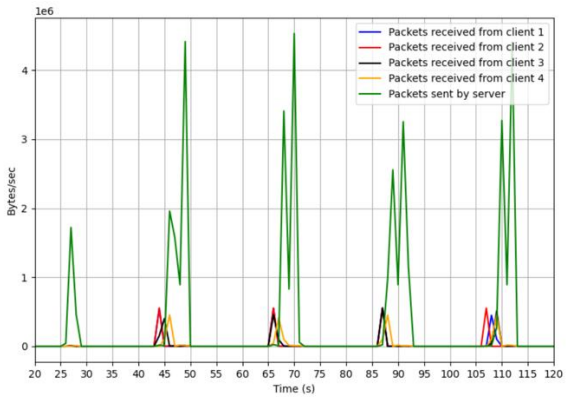


Figure 3-8: FLOWER framework. 4 clients and 1 server, clients 1, 2, 3, with 50 ms bidirectional delay, client 4 with 150 ms bidirectional delay, and bandwidth limit 1 Gbit/s

As before the graph represents the traffic captured on the server and in particular, the green line shows the traffic outgoing from the server while all the other lines refer to the incoming traffic sent by the clients. As expected, the client with the higher delay is the bottleneck of the whole network, because when the other clients finish their data fitting task and send the parameters to the server, the latter has to wait for all the clients' results before aggregating the parameters and starting a new round. In terms of accuracy, overall IBMFL was between 0.96 and 0.97, while FLOWER reached 0.98.

The takeaway message is that accurate, as well as efficient operations of FL algorithms in a distributed environment, such as an urban IoT one, are likely to be affected by latency and bandwidth conditions. Therefore, an appropriate resource allocation and provisioning strategy for communication resource must be combined with FL techniques, a direction which will be further investigated in the prosecution of the project.

3.5 Ambient backscatter communication: recycling radio waves

In this section, we consider a new method of using radio resources by recycling radio waves based on the **new and recent concept for a sustainable IoT: the Crowd-Detectable Zero-Energy-Device (CD-ZED)** described hereafter.

A new type of device known as a "Crowd-Detectable Zero-Energy-Device" (CD-ZED) has been introduced very recently [PBR+21] [Ora21]. A CD-ZED is similar to a Radio Frequency Identification (RFID) tag: it is energy-autonomous and it backscatters radio-waves to communicate. However, compared to the RFID tag, the CD-ZED has the following differences: 1) it harvests ambient energy such as solar or indoor light to power itself; 2) it backscatters ambient waves from 6G devices to be detected by 6G base stations, and vice versa, it backscatters 6G base stations waves to be detected by 6G devices, as long as the CD-ZED is close proximity to 6G devices. As a consequence, compared to the RFID tag, the CD-ZED has the following advantages in terms of sustainability: 1) it is detectable on national coverage of a mobile network, 2) it does not require the deployment of dedicated RFID readers and portals, 3) it does not require the emission of additional radio waves. The participation of the 6G device to such crowd-detection is obtained through the consent of the user. In exchange of this consent, the user can benefit him/herself from CD-ZED based services or be rewarded with some customer advantages such as lower prices on new services (after a threshold of number of crowd-detections to which the user as participated, has been reached for instance).

A sustainable IoT use case: "Smart Tracking out-of-thin air"

The CD-ZED has already been identified to provide "smart tracking out-of-thin air" services (i.e., tracking without additional energy, additional radio-waves and additional network equipment and devices) [PBR+21]. More precisely, a package under the responsibility of a logistic/transport company, bears a CD-ZED, and each time the package gets close to a smartphone (that is connected to the network and geo-located by the network), it is automatically detected by, time-stamped and localised by the network. The crowd of smartphones participates in an anonymous manner to the tracking of the CD-ZED in areas where no RFID portals are deployed.

Initial experimental tests and trials on the field of "Smart Tracking out-of-thin air" use cases

In Hexa-X, we have conducted tests and trials on the field with an improved version of the experimental set-up presented in [PBR+21], comprising a CD-ZED, a Software Defined Radio (SDR) reader emulating a smartphone. In these first experiments, our current SDR reader, implements a non-coherent power threshold detector, which is very sensitive to the level of data traffic. As the cellular network data traffic is very bursty, for these first experiments, our CD-ZED backscatters a more stable ambient commercial radio source, that is very close in frequency range to the low band 4G bands at 700 MHz: the TV signals at 642 MHz. For various measurements of the ZED-to-reader communications, we have reported three measured data:

- The maximum ZED-to-reader distance d, with a successful ZED-to-reader communication.
- The positioning error D, due to Global Positioning System (GPS) (i.e., the distance between the true coordinate of the measurement and the coordinate determined by the GPS of a smartphone).
- The ZED additional positioning error, d/D due to the ZED -to-reader distance.

The locations of the measurements are reported in Figure 3-9 a) and few environments are illustrated in Figure 3-9 b). In all locations, the ambient source (the Eiffel Tower in Paris at around 6 km) was always in non-line-of-sight. Also, in most locations, GPS positioning was challenged, either by indoor or surrounding 4 to 6th floor-buildings.

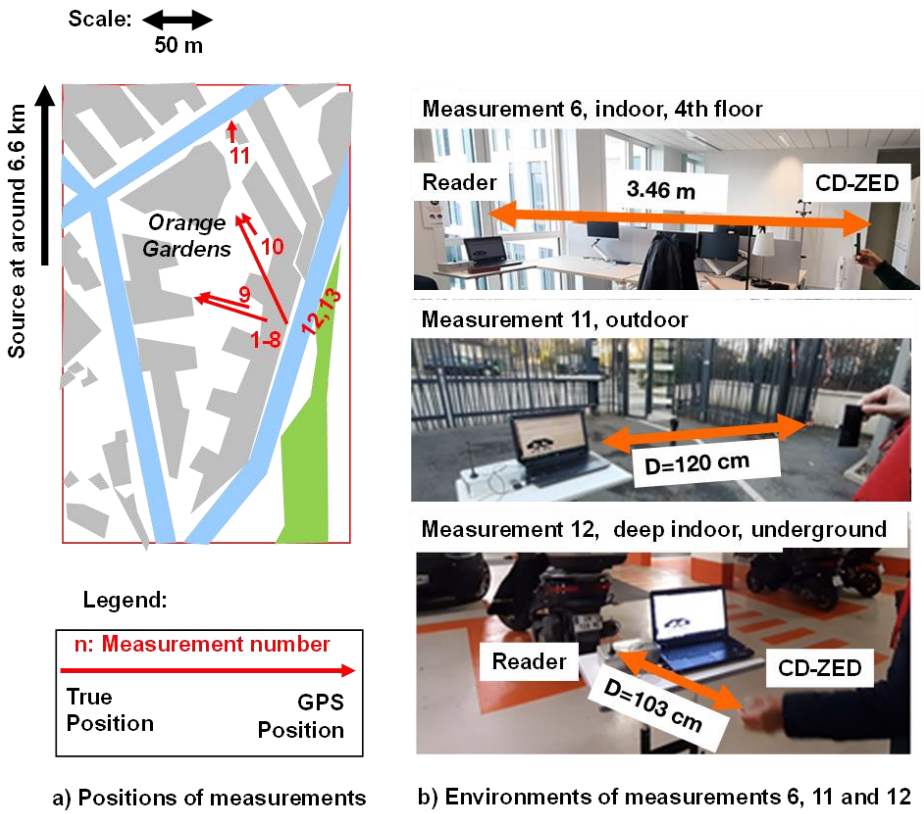


Figure 3-9: Measurement setup.

Table 8, shown below, reports the GPS positioning error and the additional positioning error (100*d/D) in %, due to the ZED -to-reader distance. As illustrated in the table, whatever the environment (indoor, outdoor, underground, etc.) the additional positioning error does not exceed 5%.

Table 8: Position error in meters, 642 MHz.

Position	Environment	Height	Positioning error due to GPS D (meters)	ZED -To-Reader Distance d (meters)	Additional Positioning Error (100*d/D) (%)
1	Indoor 6-Floor Building	Below Street	~60 meters	1.30	~ 2
2		Ground Floor		0.78	~1
3		1st Floor		0.84	~1
4		2nd Floor		1.43	~2
5		3rd Floor		0.97	~2
6		4th Floor		3.46	~6
7		5th Floor		2.16	~4
8		6th Floor		2.30	~4
9	Outdoor Surrounded By Buildings	Street Level	~20 meters	0.98	~5
10			~23 meters	0.94	~4
11			~10 meters	1.20	~12
12	Indoor	Underground Parking Lot	~176 meters	1.03	~1

Conclusions and next steps

Initial measurements show that the precision of the positioning of the ZED is similar to the precision of the positioning of the smartphone, as the additional positioning error due to the ZED-to-reader distance is negligeable. Next trials will exploit a coherent detector instead of a non-coherent detector and 4G ambient waves instead of ambient TV waves. Results will be reported in D7.3.

3.6 Flexible adaptation of functional splits in 6G RAN

6G faces new challenges to meet the needs of Internet-of-Things and Industry 4.0 environments connecting human, physical, and digital worlds. 6G RAN shall provide efficient spatial and temporal wireless network availability in line e.g., with sustained data throughput or maximum low latency. For 6G, KPIs dynamically vary over time, area and over diverse new applications and extremely high range. To efficiently address these challenges there will not be a one-size-fits-all radio access network (RAN). A flexible, adaptable and scalable RAN functionality implemented in hardware and software is needed.

The 3GPP technical report TR 38.801 [38.801] already introduced a first approach of creating and defining a multi-split RAN architecture for 5G. The methodologies of network function softwarization and RAN disaggregation are approved for 5G networks, since RAN disaggregation optimizes the total cost of ownership (TCO) and drives flexibility. Disaggregation of 5G RAN involving functional decomposition, multi-vendor operation over open interfaces and decoupling of software (SW) and hardware (HW) enables coordinated transmission and reception [GIC+18] [HR18]. Disaggregated RAN better address the increased inter-cell interference that dense mobile network deployments entail and leverages multiplex gain by resource pooling. In order to leverage the advantages functional decomposition, 5G RAN functions are split into a distributed unit (DU) and a centralized unit (CU), as illustrated in Figure 3-10 [HSC19], [SAM21].

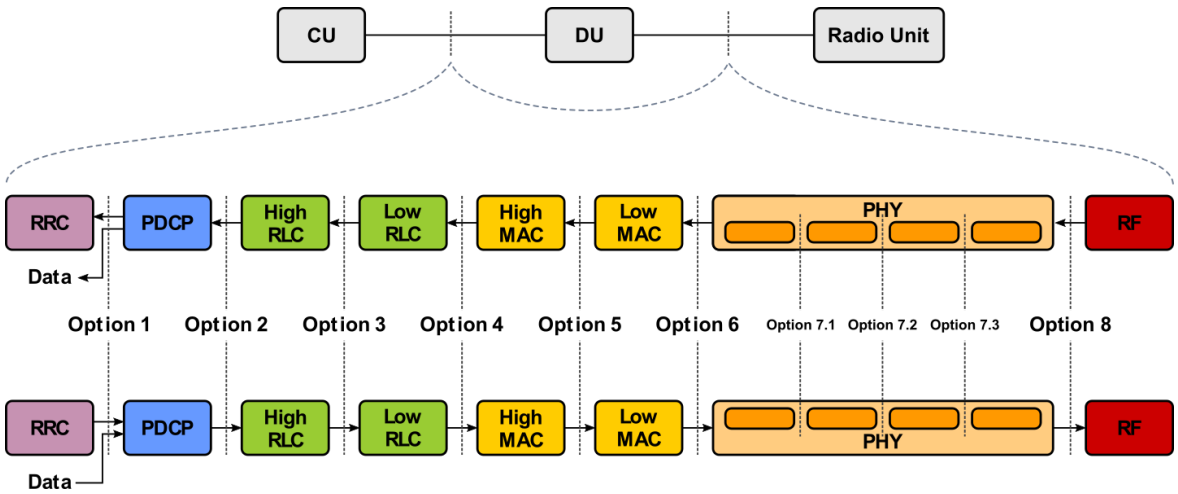


Figure 3-10: RAN functional split options for a 5G protocol stack. The upper chain represents the uplink, whereas the lower chain represents the downlink.

The current status of 5G RAN disaggregation and decoupling of SW and HW will not be sufficient for the ambitious requirements of 6G networks. The static or semi-static 5G RAN disaggregation of network function does not track the rapid KPI evolution of e.g., network availability, reliability, bounded maximum low latency or sustainable maximum data throughput that 6G networks must provide.

Use cases such as *"immersive smart cities"* envision a huge number of devices requesting sustainable maximum data throughput, bounded maximum low latency connections while exhibiting large temporal and spatial variance of these required KPIs. Other use cases such as "*autonomous supply chains"* or

"*collaborative robots"* or *"immersive telepresence for enhanced interactions"* [Hexa-X D7.1] impose high data throughput and high-availability and reliability requirements for potentially short periods of time followed by long inactivity periods. So even if a certain RAN functional split option is optimal at a given time for a given use case, the dynamic nature of 6G networks quickly renders it inefficient when compared to alternative ones. Changes in the downlink and uplink traffic, user mobility, and the appearance of users requiring special connectivity heavily influence the selection of the optimal RAN functional split option. A flexible RAN functional split that is able to dynamically adapt the underlying computing and communication architecture and HW + SW resource allocation of its RAN functions, according to the instantaneous network needs, is strongly required.

A centralized RAN is not optimal when pursuing ultra-low latency communications. The enhanced interference reduction enabled by RAN function centralization is beneficial, but low-latency use cases rely on RAN functions operating at the network edge. For those use cases a disaggregated RAN, implementing a major part of the computing and communication at edge units, is needed. A dynamically adapting functional split in a 6G RAN substantially determines 6G network KPIs for considered use cases and is briefly summarized as follows:

- **RAN reliability:** as with user throughput, by dynamically countering inter-cell interference the reliability of the communications can be optimized, allowing novel use cases such as telepresence robotics or digital twins in manufacturing. The reason is that, when Medium Access Control (MAC) or Physical (PHY) layers belonging to different gNodeBs are deployed at the CU (Options 6 to 8 as shown in Figure 3-10), it is possible to coordinate their operation and the transmission and reception of radio signals. This allows operators to prevent inter-cell interference by leveraging this coordination. Nonetheless, these splits require high-capacity links at the midhaul and fronthaul networks, since low splits entail the transmission of wideband low-level signals. If not managed correctly, changes in user traffic patterns or requirements might lead to network congestion, thus negatively affecting RAN reliability.
- **Bounded maximum low latency:** the impact of a flexible adaptation of a 6G RAN functional split is twofold. On the one hand, the improved interference management of function centralization leads to fewer packet losses and retransmissions, thus lowering the overall latency. On the other hand, for ultra-low latency applications, the network operator may prefer to move the network intelligence to the extreme edge, that is, to the DU/CU. In that case, lower latency may be accomplished by rejecting the function centralization featured in functional splits such as Option 6 and higher and instead opting for deploying all network functions close to low-latency applications at the network edge (Options 1 to 5 for the CU/DU split).
- **Sustainable maximum data throughput:** centralized functions are able to coordinate their transmissions and receptions so as to reduce intercell interference and thus increase data rates. Nonetheless, the capacity of the transport network connecting DUs and CUs as well as power and resource constraints limit the application of function centralization. The instantaneous activity and requirements of the users, as well as its geographical location, has to be taken into account when selecting which functions should be centralized. As a result, owing to the same reasons as with RAN reliability, implementing Options 6 to 8 for the CU/DU split would lead to higher sustainable data throughputs with respect to other options, at the expense of increased midhaul network usage.
- **Total cost of ownership:** function centralization leads to lower operating cost owing to resource pooling, which ideally could also reduce deployment costs. However, static function centralization is frequently not feasible, and thus the additional costs of managing and orchestrating a dynamically adapting networks must be taken into account. Recent work on this issue [AJK21] concludes that, indeed, the ability to dynamically change the functional split must be carefully adapted to the expected network behaviour, which may result in substantial cost reductions. Consequently, high function distribution, such as in Options 1 or 2 for the split, may lead to higher processing expenses since distributed unit and radio unit operation is more costly than that of the CU. On the other hand, when many functions are centralized, such as in Options 6 to 8, the cost associated with running the midhaul network may be higher.

4 Dependability in I4.0 environments

This section reports the advancements in the modelling, evaluation, and applications of the 6G dependability concept in future I4.0 environments.

In Hexa-X, dependability is defined as the "ability to perform as and when required" and identified as a critical KPI area that contains numerous related KPIs, including but not limited to availability, reliability, safety, integrity, and recoverability. These underlying KPIs serve as an indicator for Hexa-X's key value trustworthiness. In the following, we describe intermediate efforts and results, which are mainly involved with four related areas. In Section 4.1, the communication-computation-control-codesign of networked industrial systems is investigated, which evolves around the fundamental idea of converging (wireless) communication and industrial application. The identification of errors in robotics scenarios as essential component to deal with faulty behaviour on the application level is investigated in section 4.2. In section 4.3, the dependability of massive MIMO connections in investigated. This is of particular importance because the error causes in the high frequency range significantly differ from those of sub-6-GHz. In section 4.4, network-aware digital twinning is analysed as novel tool for efficient radio resource management. Therein, the (known) trajectory of mobile agents is included in network resource allocation and planning to ensure ultra-reliable connections. In section 4.5, preliminary results of the quantification and monitoring of end-to-end dependability are presented. End-to-end guarantees are crucial for industrial manufacturers, which also extends to the control plane of programmable networks. In addition, a framework to monitor dependability in a current 5G system is presented.

4.1 Communication-Computation-Control-Codesign

4.1.1 Motivation

Communication-Computation-Control-Codesign means a convergence of the building blocks that enable distributed, scalable, and dependable systems. This can be achieved through an improved understanding of how applications behave under real-world constraints such as energy limitations, Electromagnetic Field (EMF) limits, packet losses, etc. Traditional wireless communication is not designed to achieve such goals as they typically follow a "one-size-fits-all" approach. Understanding the reaction of applications to communication and computation imperfections enables an application-centred dependability evaluation, which is crucial for industrial environments. Multiple research directions were followed in the reporting period and are detailed in the following.

4.1.2 Identifying the Bottleneck

The aim of this section is to show how a joint design of communication and computing is fundamental to identify the bottleneck of connect-compute service and networks, taking into account different performance indicators and sustainability targets. Specifically, the focus is on computation offloading services, whose aim is to transfer data from end devices (e.g., sensors in industrial environments) to nearby Edge Servers (ES) that run computationally heavy applications on their behalf. In such connect-compute services, the overall delay comprises communication (to upload data) and computation (to process data) delays, thus calling for a joint design and management of these resources, to achieve the best trade-off between different performance metrics. A detailed gap analysis can be found in [Hexa-X-D7.1]. In this study, these performance metrics are: i) The offloading data rate, i.e., the number of bits offloaded per unit time; ii) The E2E delay, comprising communication and computing; iii) The energy consumption of end devices, due to uplink transmission; iv) The energy consumption of the edge server, due to computation; v) The EMF exposure in selected areas within the one covered by the service. The reference scenario is depicted in Figure 4-1, with a sensor network offloading data, and selected areas (e.g., the ones in red in the figure), in which exposure limits are required to be below predefined thresholds. Also, a queueing system is represented, to take into account the buffering delay at both communication and computation sides.

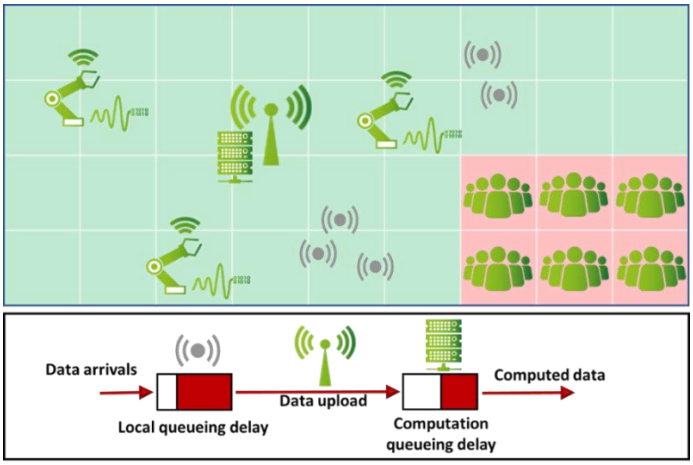
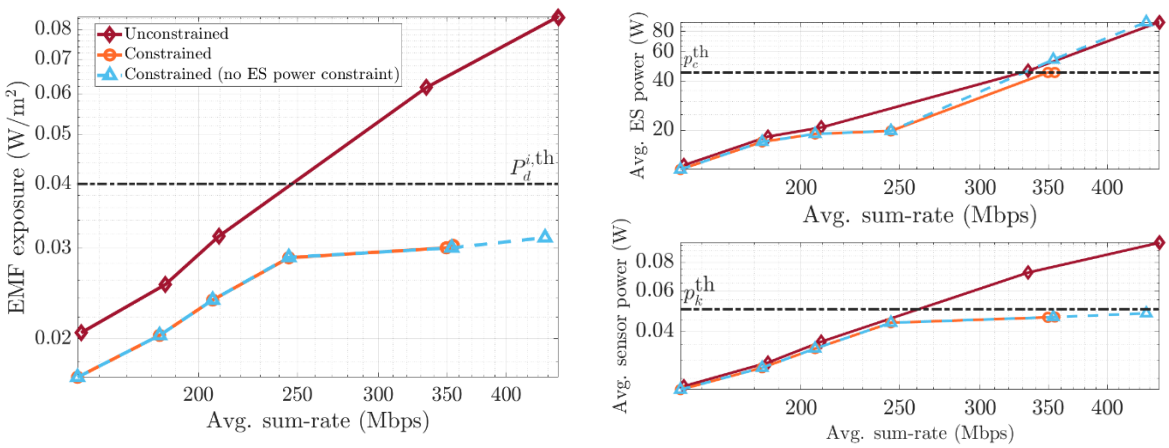


Figure 4-1: Reference offloading service scenario: End devices (sensors or machines carrying them) upload data to a server co-located with a radio access point, with target power consumption, EMF exposure, maximizing the offloading sum-rate under delay constraints.

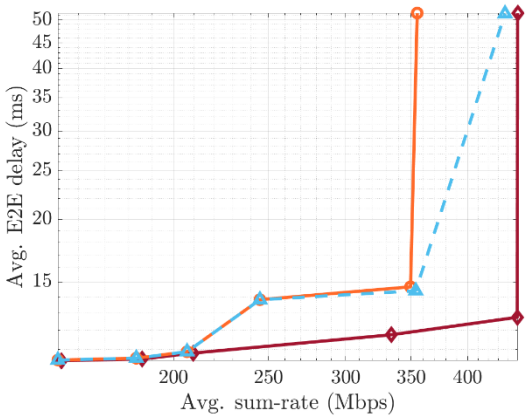
The goal of the optimization, involving radio parameters (devices transmit power and data arrivals), and computation parameters (edge server CPU scheduling) is to maximize the offloading sum-rate under delay constraints and constraints on devices and edge server power consumption, while guaranteeing EMF exposure limits in the selected zones. To this end, an online resource allocation algorithm is proposed, which relates this work with the one presented in Section 3 on flexible resource allocation. Here, radio and computing resources are dynamically and jointly optimized to achieve the objectives. More technical details on problem formulation and solution can be found in [MBC22].

Numerical evaluation

For the numerical simulations, an AP located at the centre of a squared area of side 15 m, operating at carrier frequency 3.5 GHz, and with a total available bandwidth of 20 MHz, is considered. Around the AP, in the same area, we consider 100 sensors uniformly randomly distributed in space, and transmitting with a maximum transmit power 0.1 W. Also, we assume a noise power spectral density N0 = −174 dBm/Hz. The path loss model is taken from [38.901], considering a dense factory plant. A Rayleigh fading with unit variance and coherence time equal to the slot duration of 5 ms is assumed. A single edge server is collocated at the AP, operating with a maximum CPU clock frequency of 4.5 GHz, tuneable to save energy. The different benchmarks, powers and EMF exposure thresholds are specified later.



(a) (b)



(c)

Figure 4-2: Trade-off between offloading sum-rate, EMF exposure (a), system power consumption (b) and E2E service delay (c).

As numerical results, we show the trade-off between sum-rate, EMF exposure, power consumption, and the E2E delay of the offloading service. We present three cases: i) **Unconstrained**: No limits on power and EMF are imposed, i.e., this is the solution achieving the highest possible offloading sum-rate; ii) **Constrained**: Constraints on EMF exposure over the space and power consumption (including sensors and ES) are imposed; iii) **Constrained (no ES power constraint)**: Constraints on EMF exposure over the space and power consumption (including only sensors) are imposed.

Let us show the results in Figure 4-2, first with focus on the unconstrained setting, represented by the red curves in all plots. Our aim is to show the EMF exposure, the system power consumption, and the E2E delay, as functions of the average sum-rate. Note that the optimization strategy allows the network to reach the optimal solution in terms of average sum-rate. However, approaching the optimal solution is paid in terms of average EMF exposure, power consumption, and E2E delay, as visible from Figure 4-2. In particular, Figure 4-2 (a) shows the EMF exposure as a function of the average sum-rate. We can notice how, by increasing the average sum-rate, the EMF exposure increases as expected in the unconstrained setting, up to a maximum value around 80 mW/m^2, corresponding to the maximum average sum-rate (around 450 Mbps). At the same time, in Figure 4-2 (b), we show the average device and ES power consumption as a function of the average sum-rate. Again, by increasing the sum-rate, the average power consumption of both devices and ES shows an increasing behaviour towards its maximum value (100 mW and 90 W, respectively), thus also increasing the offloading sum-rate towards the maximum value.

On the other hand, if we look at the constrained case, represented by the orange curve in all plots, we can notice how the method is able to guarantee all the required constraints (thresholds are represented by the horizontal black lines), in terms of exposure and power consumption. Moreover, EMF and sensor power consumption values are strictly lower than the thresholds, which at first glance may suggest a sub-optimality of the solution. However, the ES power consumption approaches the desired threshold, suggesting that the system sum-rate (around 310 Mbps in this case) is limited by the ES (i.e., the bottleneck is the computing part) in this setting, and the power consumption of the sensors cannot reach higher value since the ES cannot accept a higher data rate while guaranteeing system stability, i.e., finite E2E delay.

To validate this remark, we run the same simulation, removing the ES power constraint, and achieving a result represented by the light blue curve in all plots. As expected, since the system is limited by the ES capacity in the previous case, by removing this constraint, the method is able to guarantee sensor power consumption and EMF constraints, without degrading the performance in terms of sum-rate and

E2E delay, as visible in Figure 4-2 (c). As expected, in this case, the sensors' power consumption approaches the threshold, thus showing that, in the case the ES power is unconstrained, the limiting factor is indeed the power consumed by the sensors to upload data.

Conclusions

In this section, it has been shown how a joint management of radio and computing resources (i.e., communication-computation-codesign) is fundamental to identify the bottleneck of computation offloading services, which may reside in the communication or the computation part in connect-compute networks and services. First of all, an online method to manage radio and computing resources is beneficial to explore the best trade-off involving several key performance and value indicators. Moreover, the identification of the bottleneck is the first step for system design considerations that span from architectural aspects to online system operations.

4.1.3 Negative Packet Loss Correlation

The joint modelling of control and (erroneous) communication provides crucial insight for a codesign that targets to ensure dependable application behaviour and at the same time reduce wireless network resource consumption. It is well-known that operating a closed control loop requires timely data, however, the rate of transmission is another optimization target to consider in scalable industrial settings.

Two approaches have been investigated. The first is termed "reduced sampling", which provides a valuable first assessment of the control cycle's reaction with respect to packet losses. It is ultimately simple to apply, however, it is based on the conservative assumption that packet losses occur cyclically (compare Figure 4-3). This simple analysis relates the sampling period of the application with its packet loss tolerance. Established digital control design guidelines suggest oversampling the application by 10x-30x [FWP97], which introduces an inherent packet loss tolerance (otherwise it would not be oversampling). Through the reduced sampling method, this tolerance can be quantified in terms of standard LTI control theoretical metrics, e.g., the phase margin. Especially short packet loss sequences do not have a significant impact on control performance.

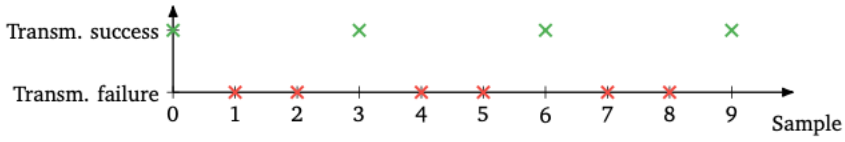


Figure 4-3: The "reduced sampling" technique assumes periodic, cyclical packet losses

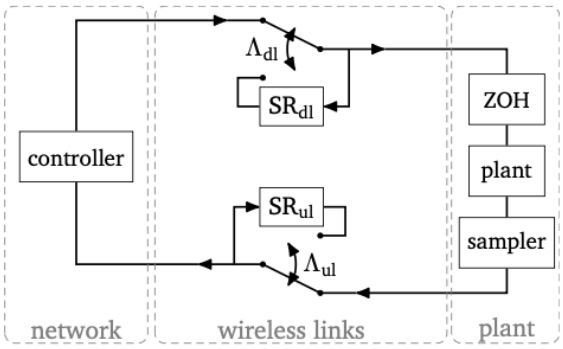


Figure 4-4: Signal reconstruction (SR) filters need to estimate values in both downlink and uplink if packets are lost

The second more sophisticated approach is based on Markov Jump Linear System (MJLS) theory. It enables deriving theoretical stability boundaries for linear-time-variant (LTV) control systems that can be described as a set of switching linear-time-variant (LTI) control systems as long as the switching

behaviour can be modelled through a Markov Chain. With a system model consisting of four different operating modes (UL+DL; only UL; only DL; none), the Markov Chain defines the temporal packet loss dependencies. In recent works, this Markov model was assumed to be the standard two-state model (link up or link down). However, the modelling approach here is to derive more detailed temporal packet loss dependencies and subsequently, provide theoretical stability guarantees. A temporal packet loss correlation coefficient was developed that ranges from -1 (full negative correlation: a second packet loss cannot occur after an initial loss) to 1 (full positive correlation: upon first packet loss, all packets until a maximum boundary K will be lost). As expected, the results show that a negative packet loss correlation is highly beneficial for closed-loop control applications, as demonstrated here for an example AGV use case (compare Figure 4-5). A negative correlation leads to short packet loss sequences and therefore, the results are consistent with the above "reduced sampling" method. The claim to require an operating point of packet loss rate (PLR) $<10^{-5}$ does not hold even for many demanding applications *if a negative correlation is ensured* (short packet loss sequences). It is recommended to rethink these demanding network requirements because more spectrally efficient networks may be built if the overall PLR can be set to 10^{-1} to 10^{-2} (with highly negative packet loss correlation).

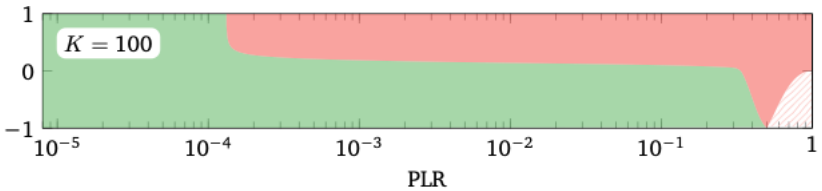


Figure 4-5: Negatively correlating packet losses (y-axis) has a significant impact on control performance. Green area = stable, red area = unstable

4.1.4 Cross-Domain Modelling of System-Level Dependability

As an inter-disciplinary research topic, Communication-Computation-Control-Codesign (CoCoCoCo) considers the impacts on the system-level dependability of all involved domains, taking the aspects of information theory, control theory and computing algorithms into account. Conventionally, the reliability problem has been studied in each of these domains, but with different focuses, simplifications, and assumptions. As an instance, from the perspective of control theory, the communication channels to send controlling commands and feedback information are generally considered error-free with constant delay. Meanwhile, when designing the communication system, the controlling loops are mostly assumed as definitely stable. In addition, the designers of computation architecture and algorithms rarely consider the existence of errors that root in the sensing process and data delivery. To thoroughly understand the system-level dependability, a cross-domain model is required to unify our state-of-the-art knowledge in these different fields, clarify the relations and interdependence among these domain-specific models.

To demonstrate the idea of cross-domain system-level dependability modelling, we carried out a case study on a simplified application of remote-controlled inverted pendulum. The system model consists of an inverted pendulum installed on a cart. The following components are integrated in the cart:

- A sensor that measures the cart position and pendulum angle
- A motor that can drive the cart along an axis in both directions
- A wireless communication module that periodically transmits sensor data to the server and receives command from it.

A model-based controlling algorithm runs on the server, which keeps moving the cart to keep the pendulum staying vertical against random perturbations. The downlink channel over which the commands are sent to the cart is assumed always reliable for simplification, while the feedback channel is lossy, over which the sensor data is transmitted back to server in uplink. If a sensor data message fails in the transmission, the controlling algorithm takes the prediction based on the last successfully received sensor data as an estimation instead. The sampling interval, controlling period, and transmission time

in both uplink and downlink are set identical. If the pendulum angle exceeds 44 degrees, a system failure is considered to occur.

As the first step, to verify the impact of controlling period on the system dependability, we set the feedback channel's packet error rate to 0 and test the system's mean time till failure (MTTF) under different sensor sampling intervals through Monte-Carlo simulations, each test lasting 60 seconds. As Figure 4-6 shows, the system remains failure free (noted in the figure as MTTF >= 60s) until the sensor's sampling interval (i.e., the feedback interval) exceeds 0.4s. Then a dramatical drop in the MTTF occurs, which is caused by the loss of controlling system's stability.

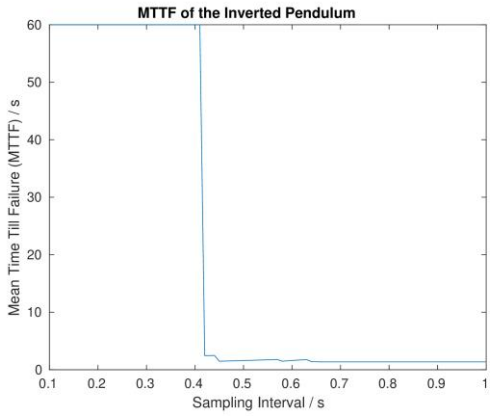


Figure 4-6: The MTTF dramatically drops to a very low level when the sampling interval becomes too long to keep the system stable, even with an ideal feedback channel

Furthermore, to verify the packet error rate (PER) of feedback channel on the system dependability, we fix the sensor's sampling period at 0.10s and measure the MTTF regarding different feedback PER. As Figure 4-7 shows, an exponential decrease of MTTF is observed, which matches the exponentially increasing state estimation error modelled in recent CoCoCo studies [HYJ+20, AVK20].

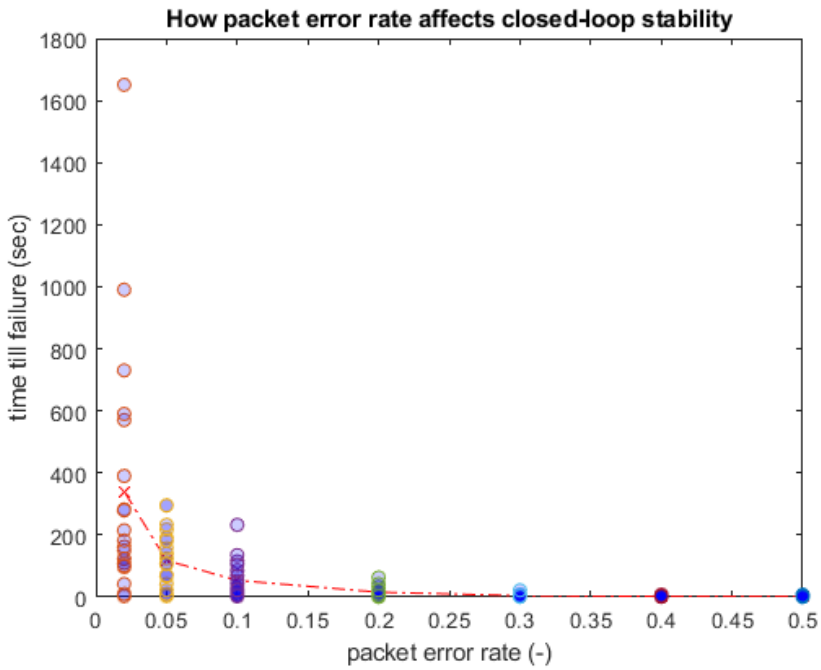


Figure 4-7: The trend of MTTF against PER when the sampling period is fixed at Ts = 0.10 s. The result of every single test is plotted as a dot, while the dashed curve illustrates the mean value.

Finally, to couple the inter-feedback interval (which is the uplink transmission time) and the feedback PER, we adopt Polyanskiy's finite blocklength error rate model [PPV10], which implies that the PER

raises along with decreasing transmission time. The Monte-Carlo simulation result, as illustrated in Figure 4-8, shows that the MTTF is a concave function of inter-feedback interval, revealing the necessity of cross-domain modelling and optimization. At the low end of inter-feedback interval, the MTTF is low due to a high uplink PER; at the high end of inter-feedback interval, the MTTF drops low again due to a lower control stability.

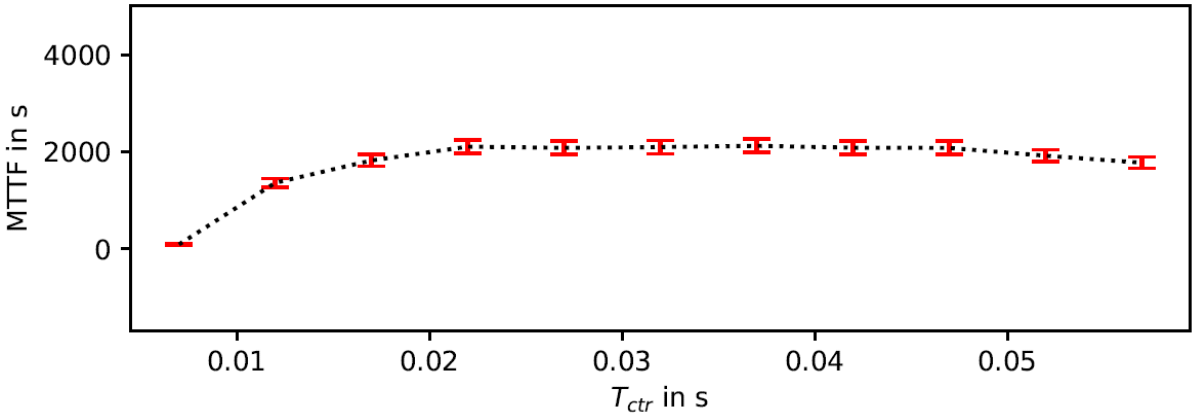


Figure 4-8: The trend of MTTF against the controlling period, which is synchronized to the sampling period. Feedback messages with finite blocklength are sent over a Gaussian channel.

Towards such a cross-domain model, there is a tech-linguistic barrier between different research fields that must be overcome first. On the one hand, the meaning of one term often diverges from one research area to another. On the other hand, different terms are also commonly used by different communities to denote the same concept. For example, when studying the feedback channel in a closed controlling loop in the communication perspective, the term "feedback delay" simply denotes the time it takes to successfully transmit the message over the channel. However, when understood in the controlling perspective, the same term indicates the overall latency between the data acquisition at sensor and the data reception at controller, which is closer to the concept of "peak age of information" in the communication field. Similarly, an "error" can indicate a failed recovery of the original data in communication, but a mismatch in value to the ground truth in computing. Moreover, the diverse understandings to a same term can still be highly related to each other, making it even more ambiguous.

For avoiding ambiguities, here we identify different error sources in the CoCoCoCo field, and cluster them into three categories, as listed in Table 9. Clearly, errors of different types have different impact on the system. While we are focusing on the system errors that directly reduces the system-level dependability, it is critical to understand how the other errors are related to system errors, and how the errors propagate in the system and eventually cause each other. To do so, it requires either an analytical approach that invokes advanced controlling theory analyses of controlling loops with random feedback delay, or a numerical solution that is driven by massive data generated from exhaustive simulation campaigns.

Table 9: Identified error concepts in CoCoCo categorized into different types

Error categories	Examples	Impact
Value errors	Sensing error, computation error, source coding noise, quantization noise, system disturbance, etc.	System state diverged from optimum
Transmission errors	Negative acknowledgement (NACK) in hybrid automatic repeat request (HARQ), packet loss	Randomly increased information delay
System errors	System outage, service failure	Reduced system-level dependability

4.2 Error Identification

4.2.1 6G network metrics supporting robotics applications

A way to ensure dependability in industrial environments is by monitoring and identifying key parameters e.g., latency, throughput, application-layer indicators, and handling possible faults in the network or the various industry components, robots, drones etc. with the minimum cost. In this subsection it is presented the network metrics and components created for error identification and treatment in I4.0 environments with 6G connectivity aiming in even lower latency and higher accuracy. These components are utilized for the PoC realization described in Sections 4.2.2 and 5.4 with the use of cooperative robots and massive twinning.

Diagnostic

Reliability, safety, and efficiency have become a priority for many technical processes and advanced methods of supervision, fault-detection and diagnosis, especially in the field of Robotics. The typical approaches of these methods, such as periodic checks of status and alerts after robotic units become inoperational, are limited and fail to give deeper insights until examined after the fact. These kinds of approaches do not allow real-time examination of the issues, and in some cases, do not even allow prediction of these events.

To overcome the above issues, a diagnostic component is being designed and developed as a specialized Performance Diagnosis tool. It is responsible for collecting information during the time that a specific application is running. Additionally, it parses that information, analyses it, and produces insights on the performance of the various elements that comprise the experiment. This component normally operates as a central collector of information, as a distinct network node or as a part of a set of management operations in a single node, retrieved from agents deployed on the individual parts of the infrastructure such as robots and other intermediate network nodes. It is also responsible for detecting anomalies in the behaviour of the various elements as well as identifying the root cause of possible problems or performance degradations during the experiment. To accommodate the real-time operation of the Functionality Allocation (FA) mechanism (see Section 3.3), the diagnostic component's operation is synchronized, to provide real-time insights as an input for the FA mechanism, when performance degradation is detected in a deployed service.

For the latter, interfaces are developed for the various internal components of the Performance Diagnosis tool to communicate with each other in real-time. This allows the processes of collection, pre-processing, and analysis of the generated data to take place on the spot, without any functional overlaps such as multiple rounds of data collection, processing and analysis of the same data from multiple components causing unnecessary traffic load on the system's networks. In addition, the Performance Diagnosis structure allows different diagnostic algorithms to be deployed. These algorithms can be tested and deployed completely interchangeably, thus exploiting the modular layout of the tool. Finally, this layout allows chaining different algorithms together for more in-depth analysis and conclusion. In this case, the results of the anomaly detection module, produced by Machine Learning (ML) algorithms, are sent to the fault localization module to determine the root of the detected anomaly.

To utilize the modular layout described above, the Performance Diagnosis tool used in Hexa-X will embody ML algorithms. Comparison tests will take place to evaluate the performance of the ML diagnostic algorithms implemented, or any other algorithms that could be deployed in the future. Root Cause Analysis algorithms are also implemented to localize a problematic node that causes performance degradation to other nodes and the deployed services.

Beside the main output of the Performance Diagnosis module, used for monitoring and diagnosing a deployed service, a secondary output is fed to the FA component's interface. The produced diagnosis is used by the FA to make decisions on a service's migration (or even termination) during the service operation, using the health status (or the root cause status, if such arises) of each service node, as

diagnosed by the Performance Diagnosis tool. An overview of the Performance Diagnosis tool's workflow is shown in Figure 4-9.

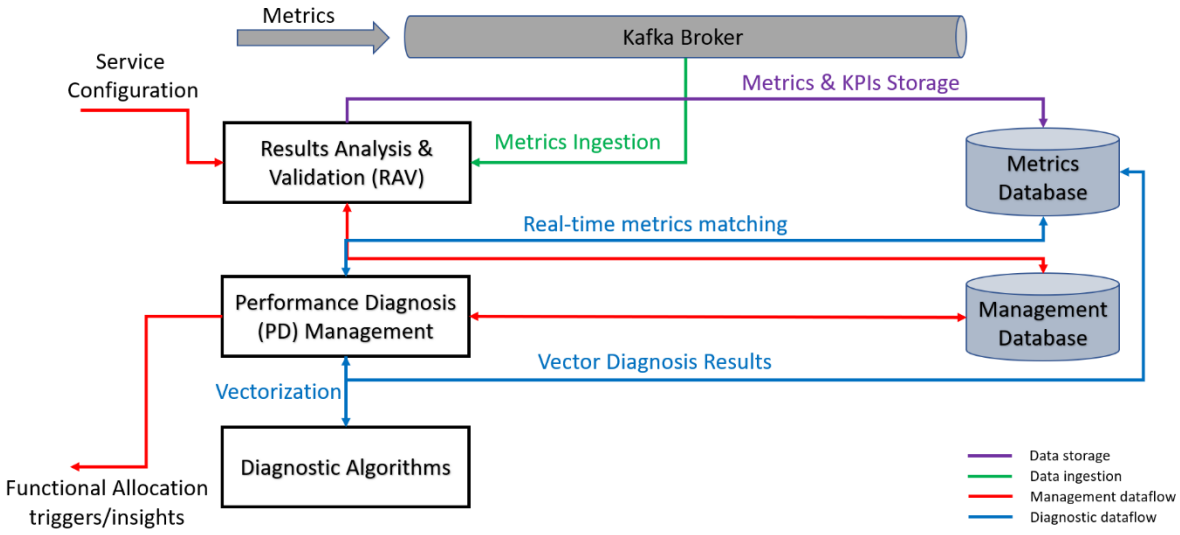


Figure 4-9: Workflow overview of the Performance Diagnosis tool.

Error identification and alerts

The error identification process is a part of the Diagnostic Algorithms sub-module and relies on the monitoring system spanning across the infrastructure. It works based on the system, network, and application metrics generated by the elements that comprise a deployed service. In this way, anomalous behaviour and performance degradation can be detected by analysis of the generated metrics. Since there is a plethora of types of errors that can occur and not all of them can be anticipated, the importance of a process that can extract information regarding the status of the various elements is required. This mechanism analyses the collected information regarding a service element from its respective metrics, logs, and other available information. According to the monitoring and metrics collection systems, a simple way to have immediate knowledge of an element's status is to relay on the status of elements in the form of a metric. In this way, when such an event is identified, the collected metrics and related information regarding the service and the element that generated the event are sent to the FA component as the context for the alert.

Monitoring reconfiguration

As previously mentioned, a secondary function of the Performance Diagnosis mechanism is informing the FA component about any alerts, which are generated either by errors or by observed performance degradations in one or more of the components. In turn, the FA utilizes AI algorithms and ingested metrics to decide if appropriate reconfiguration steps, e.g., reallocation/redeployment are needed to update the monitoring and Performance Diagnosis components, in addition to the service itself.

Typically, the workflow of the aforementioned mechanism contains the following steps:

- Error/performance degradation identification: when such an event is identified, the FA component is notified.
- FA alert trigger: all the context information pertaining to the event that triggered the alert is passed to the FA component. Additional information can also be retrieved using available interfaces.
- FA decision: the FA decides if any resource reallocation or functional placement actions is necessary.

- FA action execution: the FA decisions are executed, modifying or moving the appropriate elements/functionalities.
- Monitoring/diagnosis reconfiguration: the monitoring and Performance Diagnosis components receive the reconfiguration information to resume/restart monitoring the previously problematic element.

Functionality Allocation / Placement

The FA component described in Section 3.3, will analyse the various data coming from PD tool and the infrastructure monitor component and will decide if changes on the placement of functionalities (e.g., services, tasks) is needed for succeeding better performance or overcoming a fault/anomaly in the system.

If anomalous behaviour is detected during a service's operation, such as a node/robotic unit or a network node reporting performance degradation on a system or network level or the requested KPIs are outside the requested/acceptable ranges, Performance Diagnosis tool will detect this anomaly and alert FA component. This component will then ask for real-time infrastructure and applications information and starts analysing it. This analysis will lead to a suggestion of an optimal reallocation of the functionalities or even reallocation of more functionalities in the system, if needed. This output will then be executed by the orchestrator Open-Source Management and Orchestration (OSM) by making the necessary changes of the location of functionalities/services to system's nodes.

Accordingly, there may be a case where a node/robotic unit faces a problem (e.g., an AGV gets out of order). Thus, FA manages to reallocate all functionalities/services of that unit, so that the system can operate dependably which is of utmost importance in industrial environments among others.

4.2.2 Error identification and recovery in Robotics scenarios

In this part, we present a description of robotic scenarios with normal operation, in the case where a fault or error occurs, and the ways to handle them.

Robotic Scenario

In Hexa-X project, a use case with the title "Handling unexpected situations in industrial contexts" has been proposed. In this use case, a set of cooperative robots are working in an industrial environment (for example moving products and performing quality checks). The robots utilize modern 5G/6G networks to communicate with each other and with the infrastructure, supporting high data rate, low latency, and Human-Machine-Interface. Some of the challenges in this scenario are the cooperation of the Robots, Massive digital twinning of the whole environment, handling impairments, automated redistribution of functionality, and finally, enabling remote users to supervise and make repairs with the use of Digital Twin applications.

To showcase all the above functionalities, a specific scenario has been created. The scenario is based on two parts; 1) normal industrial operation, and 2) operation during impairment and mitigation. Three Robots (LoCoBots) are working and cooperating in a factory or warehouse environment where different automated industrial tasks are available. In addition, the user can supervise all industrial tasks, monitor the state of the robots for example, battery level, operational mode, CPU, RAM, etc. that are collected in real time and displayed at the Digital Twin application interface.

Figure 4-10 illustrates the normal operation scenario where robot R3 photographs a product and performs a quality check; meanwhile, R1 and R2 robots perform shipping and movement to repair area based on the outcome of the quality check, thus transport products from the production line to specific areas.

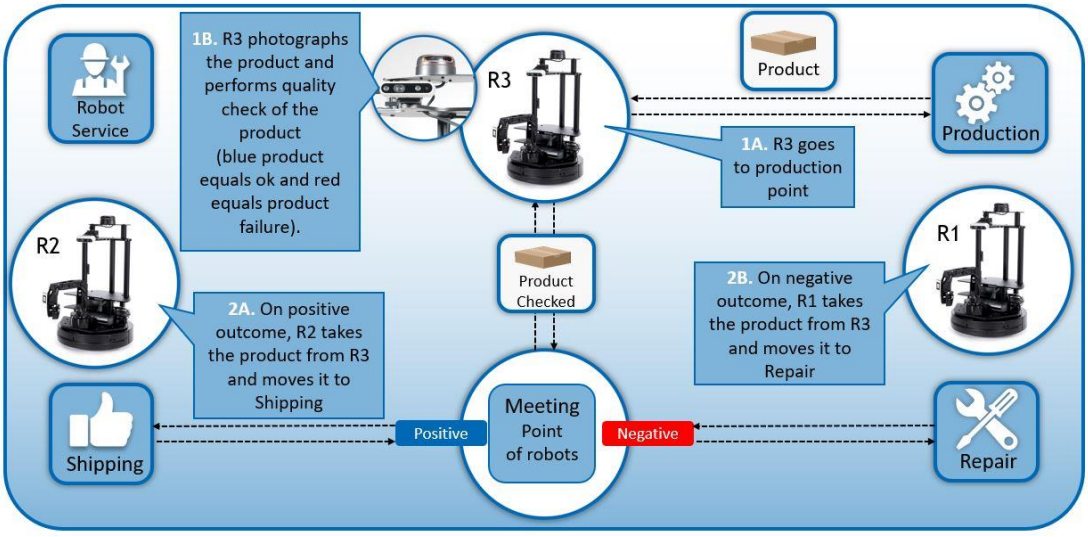


Figure 4-10: Normal operation scenario where one of the three robots performs quality check and the others transfer the product to shipping and repair area accordingly.

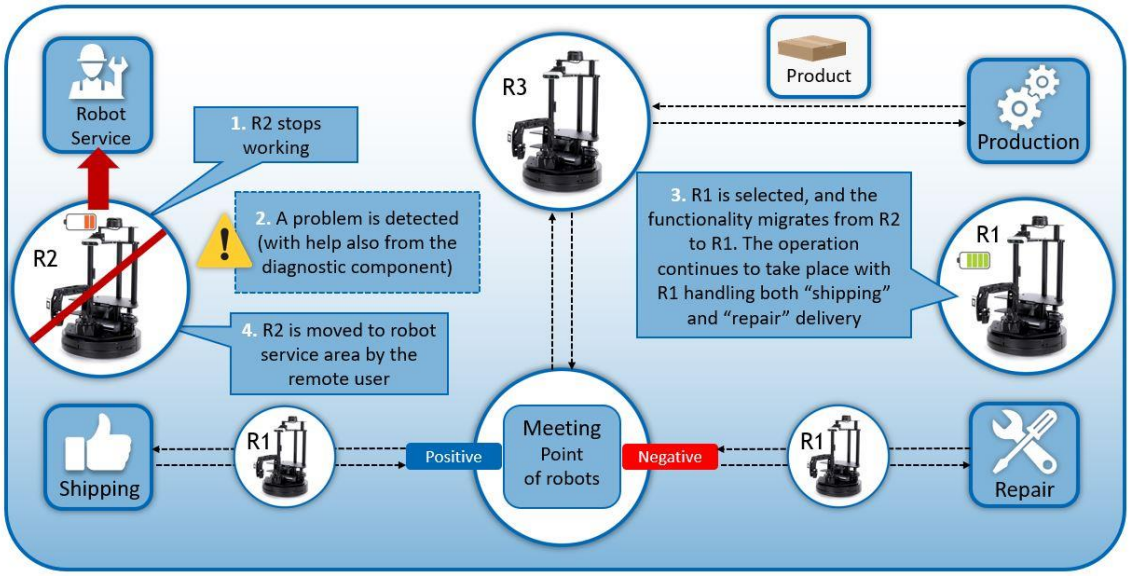


Figure 4-11: Impairment and mitigation scenario where robot R2 stops working and the functionalities are redistributed to the remaining robots while R2 is being repaired.

Figure 4-11 depicts the operation during impairment and mitigation scenario. During the normal operation, one of the robots may malfunction. For this scenario let us assume that R2 breaks down and stops working. The robot can send a notification of error, or the system can detect it with the help of diagnostic component, triggering the FA algorithm. The FA algorithm can then retrieve various data (for example CPU, memory, distance, battery level, robot-capabilities, availability, etc.) and select the best available Robot for deploying the R2's service. When a robot is selected (for example R3 Robot), the functionality will be migrated from R2 to R3, resuming to normal operation where the R3 Robot handles both "shipping" and "repair" transportation of products. In the meanwhile, the robot can move by itself or be moved by a remote user. The remote user will be notified by the Digital twin application to perform a remote repair through teleoperation, by instructing the R2 Robot to exit the production line area and go to the Robot service area to be checked and repaired by a mechanic. With the use of Digital Twin, the user can observe the operation of the robots in real time. Moreover, the user can monitor and

be informed about any problems that may occur in the industrial environment. Further, he can use teleoperation to control and solve the potential problems, thus offering a complete solution for human-machine interaction. In this way, an engineer can work from a remote location, configure, start, stop automatic tasks, interact with the robots and industrial machines in a form of a “Human-in-the-loop".

Impairment Handling

In every task, especially when multiple robots are involved, numerous unexpected events may occur. These unexpected events may be related to a network impairment, a sensor malfunction, a service failure, and many more events that were not anticipated during normal/designed operation. Thus, impairment handling and error recovery procedures are essential in order to ensure the resumption of the operation with minimal human intervention. However, the robots should be able to handle all of these cases, and depending on the failure diagnosis a different robot policy should be followed.

In case of a network impairment of a robot (see Figure 4-12), one main factor should be considered: What kind of robot services and operations were performed by the robot when the impairment occurred? On the one hand, if the resources required to perform critical tasks such as localization and obstacle avoidance were allocated to the robot, then the robot cannot continue its normal operation and should be allowed to wait for a reconnection to the network. If connection is not established within the determined period, the robot has already full knowledge of its state and environment and can safely move to a safe state. On the other hand, if the aforementioned critical resources are allocated to external infrastructures e.g., to a Multi-Access Edge Computing (MEC), then the robot should be able to recover to a safe state using a minimal set of computational resources and the input from the sensors. This recovery behaviour should be “encoded” to a service that is constantly enabled on the robot resources, but remains idle during normal operation. For the case of a mobile robot, the service should be able to recover the latest version of a local map and a significant portion of its state.

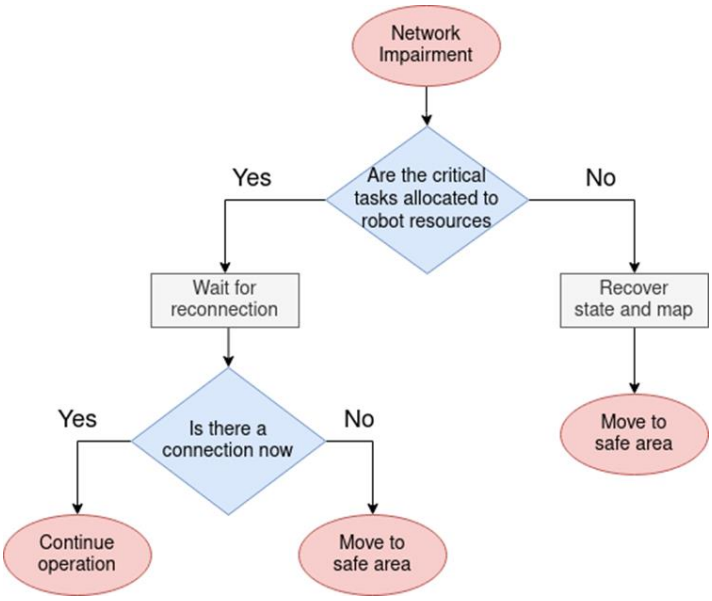


Figure 4-12: Recovery policy under network impairment.

The failure of every sensor affects differently the operations the robot is able to perform after. The most severe sensor impairment that could arise for a robot is a motor or encoder malfunction since even remote human control is unavailable. Robots usually infer them and their environment’s state using encoder values and multiple sensors such as cameras, laser scanners and IMUs (inertial measurement unit or IMU is a device that integrates multi-axes, accelerometers, gyroscopes, and other sensors to provide estimation of an objects orientation in space). If one of the latter sensors fails, then the robot should be able to return to a safe position or even complete a process cycle. If only the encoder values are left as localization information, then the state uncertainty is too high, and teleoperation is required to transfer the robot to a safe position. Finally, there are sensors that measure important scenario features

that could be allowed to fail for a shorter or larger time period such as battery level sensor, temperature sensor etc. Figure 4-13 illustrates the recovery policy in the case of a sensor impairment.

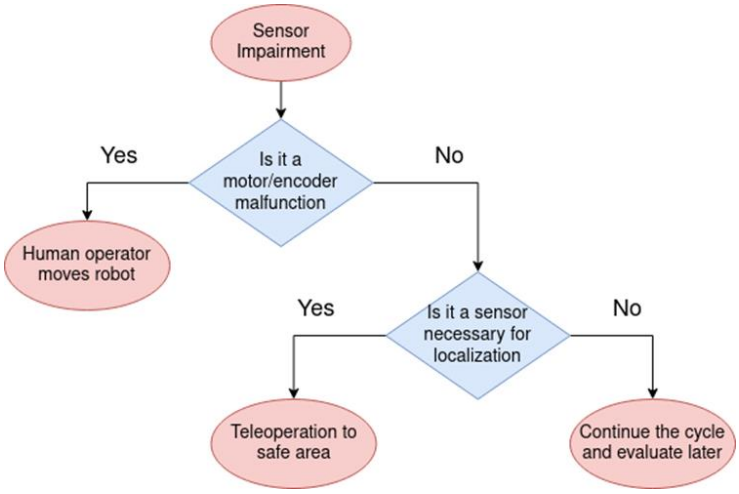


Figure 4-13: Recovery policy under sensor impairment.

A robot application is composed of multiple processes running concurrently in multiple computational infrastructures in a distributed manner. Examples of processes existing for nearly every robot are motion control, forward or inverse kinematics solving and collision detection. For AGVs, processes like localization, Simultaneous Localization and Mapping (SLAM) and image processing are essential to their normal operation. To determine the optimal recovery policy, it would be useful to distinguish processes based on whether they are stateless and on their importance for the task completion. There are processes strictly related to the autonomous navigation e.g., localization processes that require human intervention to move the robot into a safe area through teleoperation in case of a failure. These processes are considered of high importance. In addition, there are medium importance processes that allow a level of independent operation of the robot without human intervention, but they are important for finishing the cycle. Such processes could be the failure of the visual servicing process of a LoCoBot, useful for taking objects from repair or shipping. Finally, a low importance process could be a monitoring process that needs to be addressed as soon as the robot finishes its cycle. An example of operation under a process impairment is given in Figure 4-14.

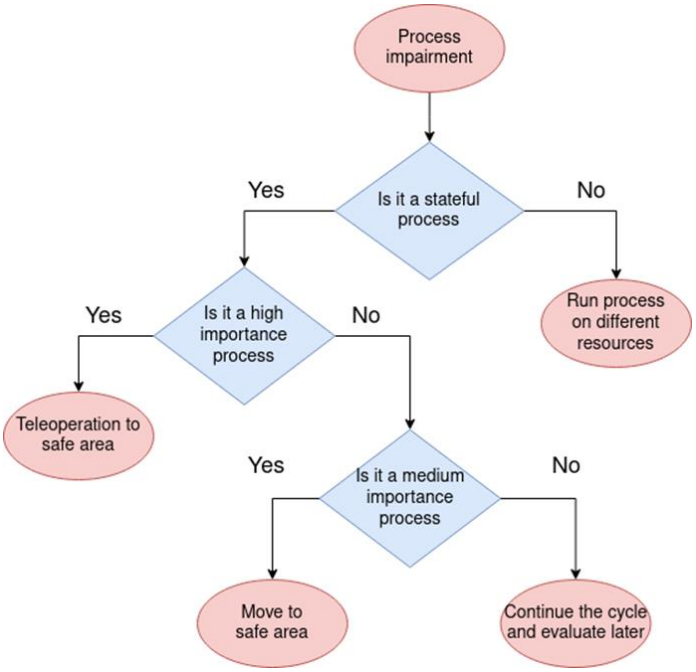


Figure 4-14: Recovery policy under process impairment.

A service failure could be defined as a higher-level unexpected event not attributed to any specific hardware or software malfunction. An example of such failure could be the inability of robot R2 to grasp the object delivered by robot R3 due to a localization error or unsuitable distance between the two robots. Unfortunately, there can be no universal online policy to handle these failures other than iterating parts of the cycle and/or replacing the robots that can be identified as responsible for the malfunction. These failures should be addressed offline when the application is originally deployed and tuned.

The return to a safe service area was probably the most common policy in case of impairment as presented above. However, when a robot is determined to be unavailable, a replacement policy should be deployed in order to achieve the highest possible throughput. Replacing the broken robot with an idle one is a straightforward solution. Also, in case of multiple idle robots, the one with the highest expected uptime should be preferred. Finally, if there are no idle robots, it is possible that one of the remaining ones could perform the abandoned task.

4.3 Dependability in Distributed Massive MIMO

The Fifth Generation (5G) of wireless communication systems introduced several communication techniques to support the coexistence of Human-Type Communication (HTC) and Machine-Type Communication (MTC) services in the same Radio Access Network (RAN). Among the novel technologies, massive MIMO was adopted in order to enhance the spectral efficiency of HTC applications that require very high data rates, e.g., mobile internet [IBN+19]. Massive MIMO consists of the use of Base Stations (BSs) equipped with a very high number of transmitting/receiving antennas, thus providing it with high spatial multiplexing capabilities. Several works have been also studying the performance gains that massive MIMO provides for MTC applications. For example, besides the high spatial multiplexing capabilities, massive MIMO also provides high Signal-to-Noise Ratio (SNR) and quasi-deterministic links, which are fundamental characteristics to ensure the dependability for critical MTC (cMTC) use cases [PSN+19]. Massive MIMO also provides high performance gains for the case of massive MTC, as shown in [LY18] and [LY18a].

Another key technique that has been considered for 5G networks is network densification, which consists of the deployment of a higher number of BSs in the same geographic area with the purpose of enhancing the spectral efficiency and number of served users [GTM+16].

However, the current wireless communication systems still rely on the cellular paradigm, where a coverage area is split into many cells, each equipped with a single base station (BS) that serves the users within the cell. As networks become increasingly denser, the inter-cell interference becomes a major performance bottleneck. Aiming at achieving the very stringent requirements of both HTC and MTC use cases in beyond-5G and 6G networks, industry and academia have already started looking for solutions for the inter-cell interference issue. The most promising solution consists of the combination of massive MIMO, network densification and a novel user-centric design: distributed massive MIMO, also known as Cell-Free (CF) massive MIMO. In a CF massive MIMO system, there are no longer cell boundaries. Many Access Points (APs) are distributed in the same coverage area, and all of them jointly cooperate to serve all the users in the area. By fronthaul connections, all the APs are connected to a common Central Processing Unit (CPU), which performs the coordination and signal processing operations [IBN+19].

A low-cost implementation scheme for distributed massive MIMO named Radio Stripes was proposed by Ericsson in 2019 [Eri21]. A Radio Stripe is a tape that has several APs composed of printed antennas and circuitry connected sequentially. The tape can be attached on any surface, thus providing spatial distribution of APs in any outdoor or indoor coverage area. The tape also provides power, time synchronization and data exchange among the APs, and connection to a CPU.

We evaluate the dependability gains (i.e., in terms of channel gain) of distributed massive MIMO networks in indoor industrial scenarios when compared to the traditional centralized massive MIMO deployment via computer simulations. We consider a square factory hall with dimensions $100\,m \times 100\,m \times 6\,m$. A total number of $Q = 64$ antenna elements can be deployed in a centralized BS or in multiple APs that are distributed in the factory hall. Each AP is equipped with S antenna elements. We name the case where $S > 1$ as a Partially Distributed (PD) scheme, and the case where $S = 1$ as a Totally Distributed (TD) scheme. Moreover, we also study two different types of distributed massive MIMO deployments: a grid deployment of APs, where the APs are deployed in a grid in the ceiling of the factory hall, and the radio stripes deployment, where a radio stripe is connected along the walls of the factory hall. Figure 4-15 illustrates the centralized massive MIMO deployments and the two aforementioned distributed massive MIMO deployments.

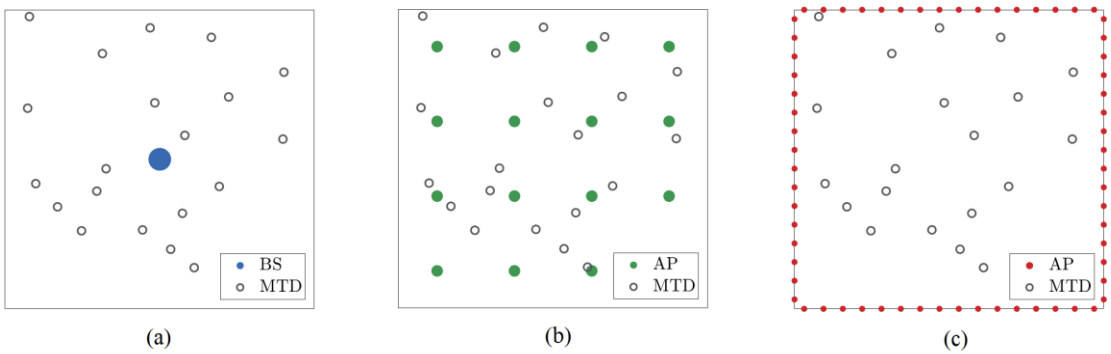


Figure 4-15: Centralized massive MIMO (a), distributed massive MIMO using the grid deployment (b) and distributed massive MIMO using Radio Stripes (c).

The goal is to evaluate the received signal strength and variability of an MTC device (MTD) in the factory hall. We adopt a large-scale fading model based on real measurements and validated by 3GPP for indoor industrial scenarios [N018]. Assuming that there is no Line of Sight (LOS) between the MTDs and the BS or APs, we also adopt a Rayleigh small scale fading model. By resorting to Monte Carlo simulations, we evaluate the empirical Complementary Cumulative Distribution Function (CCDF) of the instantaneous channel gains of a single device considering a typical case and the worst-

case position for each of the massive MIMO deployments. For the centralized massive MIMO deployment and for the grid deployment of APs, the worst-case position is any corner of the square area. For the radio stripes deployment, the worst-case position is the centre of the square area. For both typical and worse cases, the objective of our analysis is to determine which configuration offers more dependability, that is, wireless links with high channel gains and with low variability of the received signal strength. Such predictable links can be used to meet the required latency and reliability constraints of cMTC applications and also allow a more efficient utilization of the radio resources.

In order to compare the performance of the different schemes, we adopt as the performance metrics the average channel gain $E\{\|\boldsymbol{g}\|^2\}$, the standard deviation of the channel gain $\sigma_g$and the inverse of the Coefficient of Variation $1/CV = E\{\|\boldsymbol{g}\|^2\}/\sigma_g$. The average channel gain is a metric for the received signal strength, while $1/CV$ is a normalized metric for the variability of the received signal strength. In other words, a lowervalue of $1/CV$means that the wireless communication link presents a higher dependability, since it is more predictable.

The expected value, variance and CV of the channel gains for the typical case and worst-case positions are listed in Table 10 and Table 11, respectively, while the CCDFs can be observed in Figure 4-16. Both the numerical values and the CCDFs were obtained by resorting to Monte Carlo simulations. We observe that, for the typical case position, the centralized massive MIMO scheme presents the higher variability of the channel gains. The average channel gain achieved with the radio stripes setting is lower when compared to the other schemes but the channel gains present a much lower variability By using the distributed massive MIMO schemes using the grid scheme, higher values of the channel gain are achieved with higher probability because the distance between the device and the closest APs are much lower when compared to the distance between the device and the BS in the centralized case. Moreover, the variance of the channel gains, which can be noted by how steep the curve of the CCDF is, is much lower for the distributed massive MIMO schemes, which means that the received signal strength suffers from lower variations. Finally, we note that, the more distributed the antenna elements are (that is, the higher is the number of APs deployed), the better is the performance. However, such approach requires a higher fronthaul capacity since more APs should be connected to the CPU.

Table 10: Expected value, standard deviation and coefficient of variation of the channel gain for an MTD at a typical position in the factory hall and for all the massive MIMO (mMIMO) deployments.

Parameter	Centralized mMIMO	PD mMIMO	TD mMIMO	PD Radio Stripes	TD Radio Stripes
$E\{\|\boldsymbol{g}\|^2\}$	$-63.6114\ dB$	$-63.7807\ dB$	$-60.3654\ dB$	$-71.5584\ dB$	$-71.5475\ dB$
σ_g	$-57.6406\ dB$	$-60.9926\ dB$	$-56.6184\ dB$	$-68.7296\ dB$	$-70.4862\ dB$
$\frac{E\{\|\boldsymbol{g}\|^2\}}{\sigma_g}$	3.9544	1.9002	2.3697	1.9181	1.2768

Table 11: Expected value, standard deviation and coefficient of variation of the channel gain for an MTD at a worst-case position in the factory hall and for all the mMIMO deployments.

Parameter	Centralized mMIMO	PD mMIMO	TD mMIMO	PD Radio Stripes	TD Radio Stripes
$E\{\|\boldsymbol{g}\|^2\}$	$-77.5595\ dB$	$-69.8238\ dB$	$-67.2229\ dB$	$-74.4385\ dB$	$-74.4755\ dB$
σ_g	$-71.5887\ dB$	$-64.4698\ dB$	$-62.3856\ dB$	$-74.0119\ dB$	$-75.4007\ dB$
$\frac{E\{\|\boldsymbol{g}\|^2\}}{\sigma_g}$	3.9544	3.2999	3.0460	1.1032	0.8081

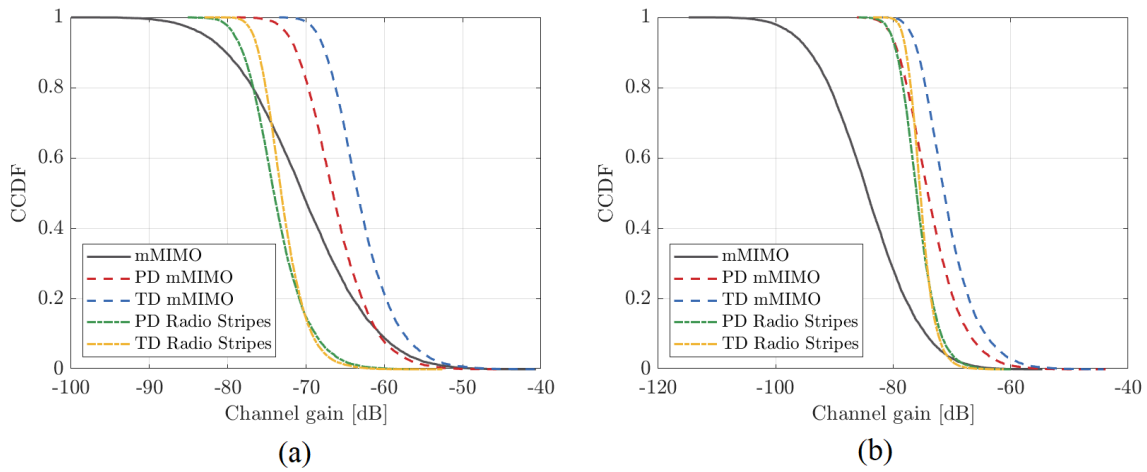


Figure 4-16: Empirical CCDF of the channel gains of a single device in the centralized and different massive MIMO deployments for the typical case (a) and worst case (b) position.

By analysing the results for the worst-case position, we attain similar conclusions. However, we note that the radio stripes setting presents much higher channel gains than the centralized setup, approaching the performance of the grid setup. Moreover, the radio stripes setting also present the lower variability of the received signal strength, which makes it the best option for use cases where high-reliability links are necessary.

Future work directions include evaluating the performance of the framework presented above for a multi-user scenario. That is, evaluating the statistical distribution of the SINR and achievable data rates considering specific requirements for cMTC and mMTC. It is also important to investigate interference mitigation mechanisms that takes advantage of the spatial distribution of APs. For example, instead of using all the APs in the network to decode the signal transmitted by all the active devices, only the subset of close APs could be used to decode the signal transmitted by an active device.

4.4 Efficient RRM with Digital Twin

In this section, we present a vision for a radio and network-aware digital twin for efficient resource management in 6G networks along with some preliminary results.

4.4.1 Introduction of digital twin technology for RAN

Future industrial environments are expected to control and optimize operations through DTs, where physical assets have a virtual/digital representation. In future factories, it is reasonable to expect that most or all the UEs are machines, e.g., AGVs, UAVs, or fixed machines. The trajectories of these UEs are known or can be controlled accurately.

A DT of the 6G RAN and core network is an up-to-date digital replica of the entire network, from the PHY to the application layer (c.f. Section 5.1). The DT maps the independent variables, namely the UE and BS locations, core network functions/nodes, radio resource assignments, transmit powers, beam selections, etc. to the KPIs of the communication services.

4.4.2 Radio-aware digital twin

In a radio-aware digital twin, in addition to the classical DT features, a set of **relevant** details of the wireless network is modelled, such as:

- The PHY and MAC layers of the access points and mobile devices in the radio network

- The properties of these transmitting/receiving radio nodes (e.g., their current position and possible pre-defined trajectories).
- Other details such as BS and UE capabilities, beam patterns, RRM algorithms etc. (queryable from RAN).
- The radio link quality in the whole network with some resolution, for example average Reference Signal Received Power (RSRP) or SINR levels on 5mx5m grid. This data can be stored (and regularly updated) in a database like a REM.

The concept is most suitable for controlled environments like private factories, and warehouses, where the environment is fully under our control; for instance, spectrum usage, UE and BS types (including their position and movement), traffic sources. The environment can also be enhanced to have better and more favourable radio conditions, for example by choosing the right materials or even by using smart materials like intelligent reflecting surfaces.

The primary benefit of Radio-Aware DT is that it enables us to predict many things in advance in an anticipatory and pro-active manner. For example, in the case of URLLC, reactively acting upon happening of errors might not be practical (when it is often already too late, for example we might not have a delay budget for retransmissions). In contrast, one can prepare for possible errors in advance. For example, we might be able to predict that a moving AGV will experience very bad quality radio link after 5 meters so we can start preparing for new beams, allocating more or different resources (even on different frequency bands) etc.

The Radio-Aware DT maintains an accurate and dynamic REM. This REM is updated constantly with a combination of enhanced measurements and advanced propagation prediction engine (e.g., ray-tracers). This update procedure makes the radio-aware DT requiring connectivity resources. In general, radio-aware DT has high situational awareness – especially in controlled environment (like private factory) where it knows the current and future positions of the UEs, and even what kind of traffic they are going to send (for example some sensors send fixed sized packets with fixed intervals). By knowing the traffic patterns, we can predict the interference patterns and estimate the SINR in a certain location. For creating the overall situational awareness, we can use many other techniques which are studied in Hexa-X, for example ML/AI (c.f. Hexa-X D4.1), high-accuracy positioning (c.f. Hexa-X D3.1) and simultaneous sensing and communication (especially with sub-THz).

We present below some early results where we have compared real-world measurements in a port environment with those obtained from an advanced physical-optics solver. The measurements were performed with two columns of shipping containers 36m long placed 8m apart, with the 3D measurement setup shown in Figure 4-17. Each column had 3 containers stacked on top of each other for almost the entire length of the canyon. We used two different transmitter positions, with the first transmitter (depicted in green) placed 18.8m from the reference corner (in blue) in the X direction at a height of 23m. The second transmitter (depicted in green) was placed at a distance of 18.85m from the reference corner in the X direction and 60.5m in the Y direction, and at a height of 22m. Measurements were made in non-line-of-sight (NLOS) at 28 GHz with a channel sounder placed at predetermined points inside the canyon.

In Figure 4-18, we plot a cumulative distribution function (CDF) of the measured pathloss against the estimates obtained from a physical optics solver. It can be seen that there is excellent agreement between the measurements and the solver prediction, with an error of 1-2 dB. Further analysis of the measurement data is in progress to evaluate the difference between error between the measured parameters and those predicted by the solver.

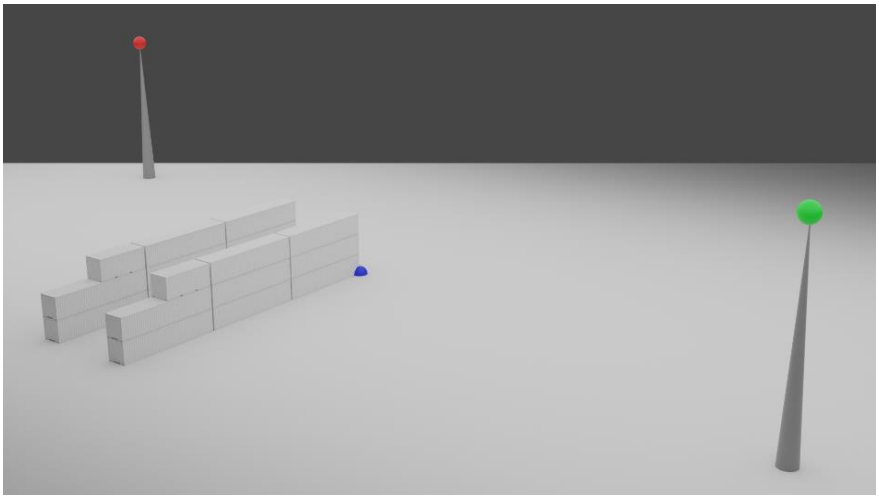


Figure 4-17: 3D model of the propagation environment.

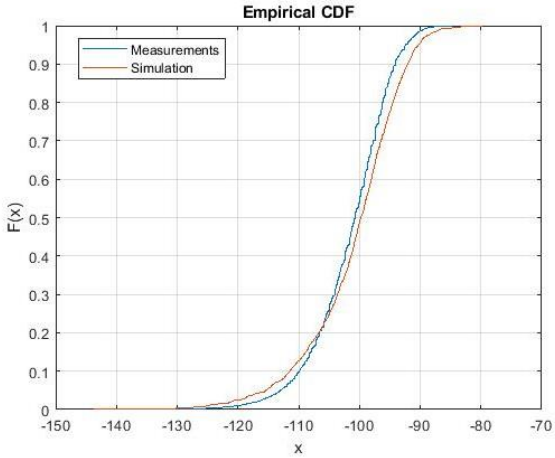


Figure 4-18: Comparison between the measured pathloss and the predictions from a physical optics solver.

4.4.3 Network aware digital twin

6G network, especially non-public or private networks (NPN) are utilized in connected process and automation industries. One of the key requirements of industry 4.0 is flexibility and reconfigurability based on the production demand. In this case, network cloudification and softwarization becomes even more prominent and favourable implementation trend due to their dynamicity and manageability properties. Thus, a network-aware DT can incorporate and virtualize the desired network resource components and performance characteristics by considering both network service user and operator's requirements for service provisioning and network operation optimization. Depending on the optimization goal, the components and characteristics in the network DT can be categorized into the two planes as follows:

- Network function/node plane. In this plane, envisioned 6G network functions (extendable from the current 5G network functions) can be modelled in network DT. Several network function/node-level parameters or attributes such as aggregated data rates or average packet delay for all served users per network function or RAN node can be included in the network DT for the usage of certain network resource or service optimization algorithms.
- Network service slice plane. In this plane, each network slice corresponds to a particular network service or application with specific QoS requirements. These slices can be modelled in the DT. The concerned characteristics of a network slice can include but not limited to, number of served UE/application packet data protocol unit (PDU) sessions, aggregated or

average UE/application session data rate or average packet delay or jitter of UE/application sessions, the percentage of UE/application sessions meeting the associated QoS requirements, and the amount of consumed radio or network resources represented by an absolute or relative metric.

To perform the intended operations, the network-aware DT is required to acquire network measurement/analytic data. Ideally, the more comprehensive and frequent the collected data is, the better performance of the DT can be achieved. Since the benefits obtained from the network -aware DT is at the cost of network data collection in terms of communication resource and energy consumption, the trade-off between the expected benefits and the cost caused by the data collection must be thoughtfully calculated. Therefore, to collect the network measurement data, two approaches can be taken, periodic and event triggered. On the one hand, with the periodic data collection, the DT can request the 6G network to provide the requested data periodically. On the other hand, the event-triggered data collection transmits the data to the DT when a particular event negotiated between the DT and the 6G network occurs. It is plausible that mechanisms to control the frequency and volume of data collection shall take into account the balance between the expected performance improvement of network-aware DT and the consequent expense of the data collection.

4.4.4 Conclusions on RRM with digital twins

A radio-aware DT (that includes the DT of the PHY and MAC layers) can help to improve the spectral and energy efficiency of a network. For instance, by having a digital replica of the network at the PHY layer, and by using powerful ray-tracing algorithms or data-driven models, it is possible to predict the channel conditions between UEs and network nodes without explicitly requiring measurements, thereby reducing the overhead in the system (improving spectral efficiency). This approach also has the potential to improve the energy efficiency since lower transmit and processing power (for e.g. sophisticated channel estimation algorithms) are spent on these measurements, provided the computational power required for the ray-tracing algorithms are lower than those required for generating and processing the measurements.

With an accurate prediction of link conditions, it is possible to put UEs on short-term "scripts" where, the radio resource management is performed for the entire T seconds and conveyed to the UEs in one shot. However, the UE trajectories over the time interval T seconds, and the estimation of the transmit buffers of both the network and the UE in this duration are needed to be known. Some of the radio parameters that are part of control signalling such as active bandwidth-part (BWP) selection, time-frequency resource allocation within a BWP, transmit power control, modulation and coding scheme can be signalled in advance to the UEs. This approach makes sense in a factory automation scenario where the environment is quasi-deterministic (we assume that the environment is fully under our control, but some aspects may not be modelled exactly), and the probability of an unprecedented event is extremely rare (much more infrequent than a packet drop).

The main benefit of putting UEs on short-term scripts is a reduction in the associated control overhead. In the absence of a radio-aware DT, the system is reactive where:

- the UE/network node performs measurements.
- if the UE performs measurements, the results are then conveyed to the network node.
- the network performs RRM based on the measurements.
- the network has to inform the UE about the resource allocation.

These steps not only increase the overall latency, but also the second and fourth steps require dedicated control signalling between the UE and the network node, increasing the overhead (also increasing energy consumption). By putting UEs on scripts, this overhead can be considerably reduced since the system is controlled proactively, rather than reactively.

By virtue of network-aware DT, it is expected that network resources are optimally managed and allocated according to the actual needs of the instantaneous services. Meanwhile, the energy consumption of different services can be efficiently dimensioned during the network design and operation phases.

4.5 Quantifying and Monitoring E2E Dependability

The need for an E2E approach to dependability was motivated in [Hexa-X D7.1] and generalized for the Hexa-X use case families and selected use cases in [Hexa-X D1.3]. Briefly summarized, E2E dependability considers dependability attributes specified for all services offered by a 6G system and consumed by an application. In [Hexa-X D1.3], the respective services are (i) communication, (ii) computation and AI, and (iii) localization and sensing. In the following, we go into more details on metrics for the individual services (some of them being generalizable) and our contributions towards increasing dependability. Lastly, we present contributions on dependability monitoring in Industry 4.0 scenarios as a fundamental building block for utilization of flexible resource assignment mechanisms and the realization of network-aware digital twins.

4.5.1 Dependability metrics and extension to E2E dependability

Dependability, security, and performance are the three factors that determine the trustworthiness of a communication network according to [SHC+10], as illustrated in Figure 4-19. Dependability consists of five attributes: reliability, availability, maintainability, integrity, and safety. There is a link between these dependability attributes and the other two factors of trustworthiness: integrity and availability also serve as attributes for security, and the QoS measures that define the desired performance are the underlying constraints described by the availability attribute.

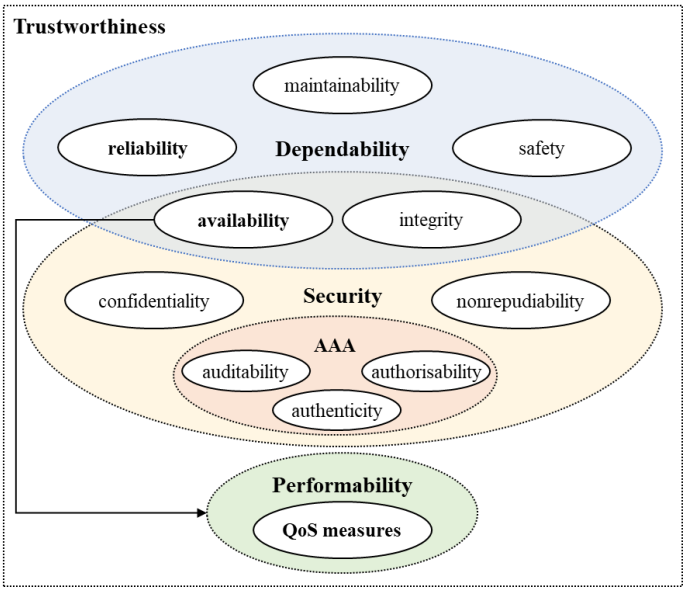


Figure 4-19: Dependability and its attributes as cornerstone of trustworthiness (adapted from [SHC+10]).

The 6G vision as described in [Hexa-X D1.2] extends the 6G system beyond communication by providing new capabilities such as computation/AI or localization/sensing to applications, requiring us to consider those additional services when describing desired QoS and dependability attributes from the application or use case perspective, as done in [Hexa-X D1.3]. We summarized the contributions towards achieving dependability for the individual services in Section 2.2.

A black box approach is no longer viable once a co-design of network and applications is considered, or the potential of the application to impact network behaviour because of self-learning and adapting network management and fine-grained AI-based optimization. Instead, we propose to follow the MAPE-K cycle [IBM06] to realize end-to-end dependability: the network and its services, its underlying infrastructure, and the application at hand need to be closely monitored (M), and the gathered data needs to be analysed (A) to allow planning (P) of mitigation or prevention mechanisms and their execution (E) in the network and/or the application (e.g., through flexible resource allocation). The whole process requires knowledge (K) to understand the implications of such measures and anticipate

application behaviour and, potentially, even human interference. This knowledge is realized by means of a network aware digital twin, as further elaborated in Section 5.6.

A significant challenge is to identify relevant indicators (events, alarms, measurements) and their correlations, as well as necessary contextual information that helps in the analysis and planning steps, as later discussed for our contribution on dependability monitoring. This process requires knowledge about the respective application and environment (e.g., specialized industrial applications and OT networks in a factory), which might include business critical or sensitive data. Therefore, it is assumed that it is realized as a local control loop outside of the scope of general network optimization and management [Hexa-X D6.1].

4.5.2 Control and data plane guarantees in programmable networks

In the scope of Industry 4.0 scenarios, achieving dependable and often deterministic communication is a key requirement, as application resilience, especially in brownfield deployments (i.e., with legacy equipment and machinery) is very limited, and QoS violations lead to either complete standstill or at least significantly decreased application productivity.

To utilize the flexibility of B5G/6G systems stemming from virtualization and programmability under these conditions, the performance of both the data and control planes of the network should be predictable. Therefore, in addition to predictably forwarding user data, the network needs to promise timely delivery of the control plane messages, and network updates and the impact of virtualization needs to be considered.

The current state-of-the-art systems either assume to have full end-host control (e.g., traffic scheduling, policing) as in data centres [AGM+10, LML+19, ZEF+15] or they focus solely on the data plane and do not provide deterministic operations at the control plane [BDZ+19, GSG+15, JSB+15]. We propose a novel system that combines predictable control plane operation into a data plane and provides joint data and control plane end-to-end latency and throughput guarantees. Our proposed system relies on an in-band control plane, i.e., without the need for physically isolated dedicated control plane channels, thus significantly decreasing the CAPEX and OPEX. Further, our system relies only on existing mechanisms available in current programmable state-of-the-art switches, such as traffic metering, priority queuing, and label-based forwarding. Therefore, it can be deployed without creating new or modifying existing protocol stacks.

The end-to-end flow requests are defined as a 5-tuple (src, dst, rate, burst, delay). These requests are arriving at the North-Bound-Interface (NBI) of the controller in an online manner. The controller uses the developed admission control algorithm to embed the flows into the network. To guarantee the performance of these flows, we must ensure that the end-to-end delay of them is bounded, and the data rate is guaranteed along the path. To achieve these guarantees, the admission control algorithm employs deterministic network calculus framework [BT01].

Moreover, improving the flow acceptance rate means more revenue for the network operator. Thus, we employ network reconfigurations to increase the number of accepted flows. In more detail, to embed a new flow, in case of rejection at the first try, we re-route some of the existing flows to free up the resources for the new flow. Thus, given the flow request *f*, the output of the admission control is the set of network updates that must be performed on the network. These network updates are either a single forwarding rule addition or a rule addition with dependant reconfigurations (rerouting). These network updates should be sent to the forwarding devices over the in-band control plane channel. To do this, we propose a fast and efficient scheduler algorithm to (re)configure the network predictably and consistently.

This procedure is designed such that the data plane traffic guarantees are not violated during the network update. To confirm it, we evaluate our system in a testbed. We use a network topology as presented in Figure 4-20. There are 5 switches, 6 hosts, and a controller in our testbed. Network switches are PICA-P3297 and EdgeCore Tofino Wedge 100BF-32X/65X.

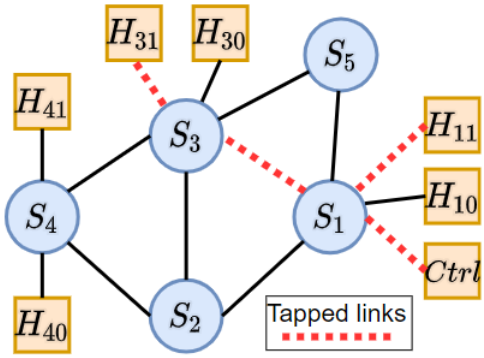


Figure 4-20. Testbed network.

To perform the evaluations, we use different sets of flow types with random source-destination pairs and various characteristics as presented in Table 12.

Table 12: The considered flow types and their characteristics (as uniform distribution).

Flow Type	Rate (Mbps)	Burst (kb)	Delay (ms)
Industrial	U(8, 40)	80	U(5, 10)
Hadoop	U(40, 120)	U(80, 120)	U(10, 100)
Data mining	U(80, 160)	U(12, 160)	U(10, 100)
Strict consistency	U(5,8)	U(80, 120)	U(5, 50)
Adaptive consistency	U(2,4)	U(80, 120)	U(5, 50)

The traffic flows are generated with DPDK-based MoonGen [EGR+15] running on a server with Ubuntu 18.04 and equipped with Intel Xeon E5-2650 v4 @2.2 GHz CPU and an Intel X520 NIC. To measure the end-to-end delay, we tap some of the links in the network (dashed lines) and measure the latency with a 10G nanosecond-precise Endace DAG 10X4-S measurement card. Next, we inject a set of random flows from the above table into the network and select a few of them to evaluate.

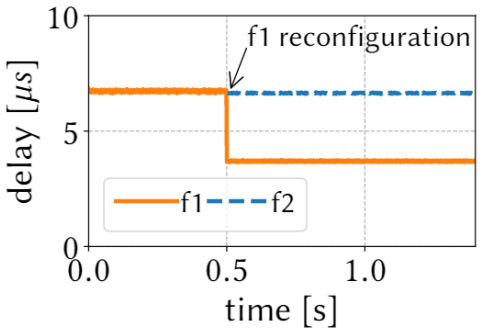


Figure 4-21. Confirming the consistency of the proposed system while reconfiguration.

Having more than 100 flows successfully embedded in the network, we did not observe any packet loss and end-to-end delay violations for all the flows. Further, to evaluate if the proposed system can reconfigure the flows without violating their end-to-end delay and packet loss guarantees. To do so, we add flow *f1* to the network at *t=0s*. Thereafter, at *t=0.5s*, we add a new flow *f2* to the same path as *f1*.

At the same time, we re-route *f1* to a new path with lower delay. We measure the e2e delay of these flows, as presented in Figure 4-21. The e2e delay remains constant during the measurement, especially at *t=0.5s* where the reconfiguration is performed. Further, during this experiment, we did not observe any packet loss.

In another experiment, we evaluate the scalability of the system's control plane, especially in terms of network update rate in a worst-case scenario. To do so, we perform a set of simulations on Internet2 topology. We vary the number of nodes in the network; hence, the number of flows. The flows are arriving at the scheduler over time. Figure 4-22 indicates that the system's control plane is scalable with the network size. Our proposed scheduling algorithm can keep the network update rate up to 280 per second, wherein the state-of-the-art approaches [ZHK+21, NCC17], this number stays around 5.

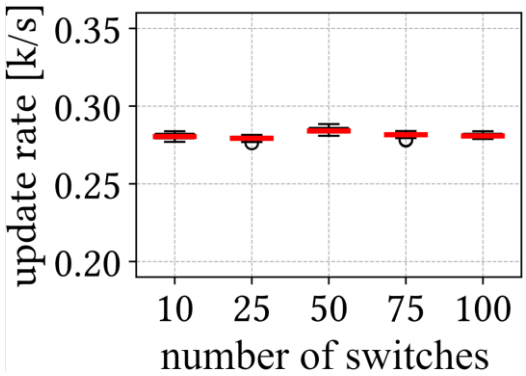


Figure 4-22. Network update rate scalability.

The proposed system acts as one building block for achieving the required QoS attributes for dependable communication while maintaining flexibility in virtualized and programmable network resources.

4.5.3 Dependability for new 6G capabilities

In addition to communication, dependability and, more generally, trustworthiness needs to be considered for additional 6G capabilities: AI/computation and localization/sensing.

Requirements for AI are discussed in detail in [Hexa-X D4.1] for two main usage scenarios: in-network learning and AI-driven air interface design. The first scenario is relevant in the scope of e2e dependability for I4.0. It constitutes a fundamental functionality for AI as a Service or Compute as a Service as new capabilities of a 6G system. In the context of Industry 4.0 use cases, certain requirements for the utilization of AI stem from standards in the respective industrial domain. One example are functional safety requirements in industrial automation, expressed as a safety integrity level (SIL). The Standardization Council Industrie 4.0 (SCI 4.0) outlines respective recommendations in [SCI20] and points to relevant standardization activities on a national, European, and international level, e.g., the committee SC 42 "Artificial Intelligence" within the ISO/IEC JTC1 joint committee.

We monitor these activities and exchange with WP4 on dependability in AI, especially in the context of the digital twin and its trustworthy execution platform discussed in Section 5. This also includes utilizing localization and sensing services, as described in [Hexa-X D3.1].

4.5.4 Dependability monitoring in I4.0 scenario

The goal of dependability monitoring is first to capture and understand the impact of certain QoS degradations on application productivity. If the application is resilient in that it reacts to specific changes and reconfigures itself to mitigate a potential impact on the overall productivity, this effect also needs to be captured to understand what QoS degradations might be tolerable in specific scenarios. This is especially relevant if Communication-Computation-Control-Codesign is applied in the respective scenario, as discussed in Section 4.1.

To this end, a testbed setup is proposed that allows continuous QoS measurements and event-based alerting across the network and its infrastructure elements and within the (virtualized) applications. The respective setup is illustrated in Figure 4-23. Infrastructure elements included in the testbed are: (i) an industrial 5G UE connected to an industrial control PC, (ii) a radio unit utilizing the 3.7-3.8 GHz band (spectrum assigned to local campus network use in Germany), (iii) an integrated DU and CU running on a single server, (iv) the 5G core running (iv-a) on a dedicated server as a native application and (iv-b) on a local self-managed two-node Kubernetes cluster.

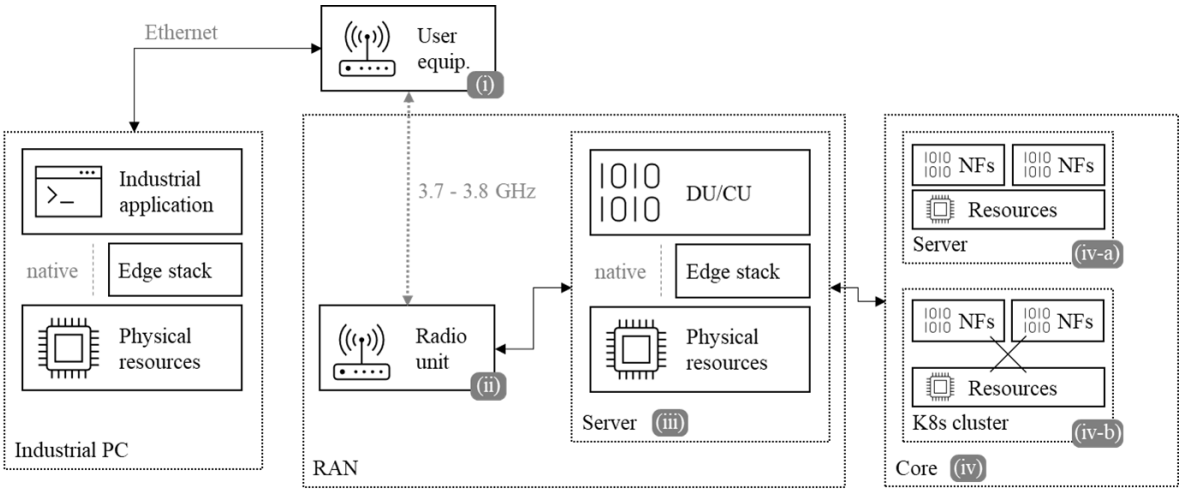


Figure 4-23: Testbed setup, with measurement data being gathered from all components.

All collected data is gathered in a time/geo-series database (here, InfluxDB is utilized), capturing the temporal sequence of events and their spatial characteristics in case of mobile components (spatial information is used in T7.2 and T7.4). Data collection utilizes readily observable interfaces wherever possible, including Prometheus REST APIs or Telegraf [5] plugins for network functions and infrastructure monitoring. In cases where such interfaces are not available, custom measurement code is utilized to capture specific logs or status information. Then, the results are inserted into the time/geo-series database using a standard monitoring agent (here, Telegraf is being used as a local agent).

For application-level QoS metrics, a "measurement container" is utilized. It contains a workload generator for traffic patterns covering specific aspects of the Industry 4.0 use case and the corresponding measurement tools. Container instance execution and control of measurements are instrumented across the testbed using Ansible, regardless of whether the container is running directly on the host OS (e.g., docker daemon) or within a Kubernetes cluster. The measurement container is an extension of the open-source Network-Multitool[6].

The gathered information is further augmented with measurement data from mobile UEs (e.g., mobile robots or AGVs, or humans wearing novel HMIs) as later discussed in Section 5.6. The resulting network-aware digital twin is aiding resource allocation and failure prevention or mitigation strategies on the control- and data plane and further allows an optimized placement of virtualized applications on available edge resources, as further elaborated in Section 3.

[5] https://www.influxdata.com/time-series-platform/telegraf/

[6] Available online: https://github.com/wbitt/Network-MultiTool

5 Digital Twins and novel HMIs

To achieve the convergence of human, digital, and physical worlds, which is envisioned by the Hexa-X project, the concept of DT and its massive deployment in the industrial 6G environment will play a key role. Combined with new concepts for HMI, DT will enable a large cluster of new use cases, which can eventually assemble into an organic ecosystem of 6G human-centric industry. In the following, new concepts, use cases, and intermediate technical solutions proposed in T7.4 are introduced. This includes an overview to the 6G industrial DT ecosystem, a model for the impact of human presence on industrial DT deployment, a solution for massive twinning with human in loop, a development of DT-empowered collaborative robots, a concept proposal of network-aware DT for local insight, and a pioneering study on exploiting DT to enable emergence intelligence in 6G.

5.1 Ecosystem of Digital Twins in Human-Centric Industrial Environment

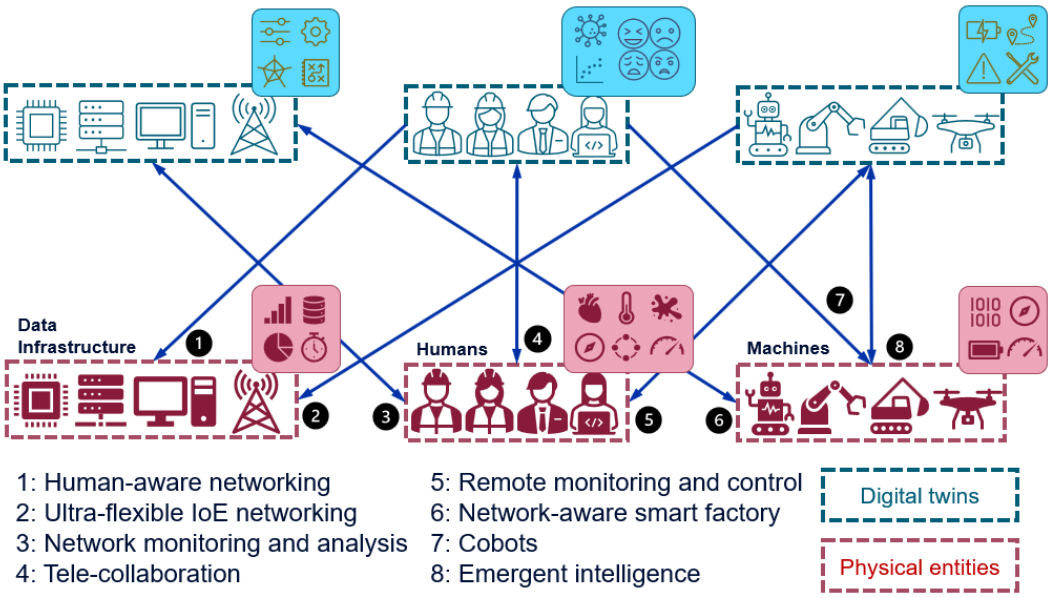


Figure 5-1. The ecosystem of 6G human-centric industrial digital twins, the arrows indicate the direction of information flow.

With the envisaged massive deployment of digital twins, a DT can be created and maintained for everything (including both the assets/equipment of vertical industry and the network infrastructure) and everyone involved in the future 6G industrial scenario. The instantaneous observable states of the physical entities are continuously measured and uploaded to the computing nodes, which are usually located at the network edge for low latency, so that the digital twins stay as up to date as possible (note that the location may also be affected by the level of trust, see more details in Section 5.6). Examples of such low-level status information are including but not limited to channel state information, computing resource availability, and user trajectory. Timely updating and archiving such data, industrial DTs are capable of not only monitoring the low-level status of massive machines and humans in real-time or almost real-time, but also estimating high-level context information that is typically unobservable, such as the physical/mental health condition of humans, reliability of industrial equipment, or the possibility of malicious cyber-attacks. Leveraging such knowledge, we will be able to develop a 6G DT-based industrial ecosystem, in which the three worlds of data, machines, and humans converge into one. With ubiquitous, fast, and reliable connectivity for everything and everyone, the interconnection of physical entities and digital twins will significantly expand the ways in which humans and things recognize and interact, enabling several revolutionary emerging use cases:

Ubiquitous and Collaborative Telepresence: The 6G network will enable bidirectional data exchange between physical entities and their DTs in a timely and reliable manner, allowing for virtual interaction with the former over the latter. Furthermore, with the ubiquitous coverage and high accuracy in time synchronization to be provided by 6G, the same digital twin can simultaneously interact with multiple individuals (or their DTs) in arbitrarily different locations without triggering conflicts or collisions between the operations. Thus, a remote collaboration that works as smooth as on-site can be realized by means of telepresence, which will dramatically reduce the efforts and costs associated with spatial distance and access barriers.

Understanding the World: Staying connected and synchronized with everything involved in the physical environment, a DT-empowered industrial system can collect a massive amount of data. Thanks to the latest achievements in the field of ML and AI, the hidden coherence and causation between different events can be extracted from the data, which will certainly enhance our comprehension to the industrial equipment as well as the physical environment, and therewith help us to appropriately design, configure, operate, diagnose, and maintain them. On top of such comprehension, complex and collaborative interactions among different entities will arise from the meaningful and selective data exchange between their digital twins.

Understanding the Humans: While industrial equipment and the physical world mostly behave in a deterministic manner that can be precisely predicted upon an accurate model, human behaviour cannot be completely predicted. Therefore, human coexistence in industrial scenarios and human participation in industrial processes introduce significant uncertainty into the system, inevitably resulting in risks of service degradation and failure. Nevertheless, with massive data collected and analysed, the DTs of human participators can create a statistical model of human behaviour that assists the industrial system in assessing and mitigating such risks. Furthermore, using multi-dimensional status information obtained via advanced human-machine interfaces, human DT can recognize higher-order human status that are difficult or impossible to observe directly (e.g., emotion or fatigue). Numerous additional benefits can be therewith obtained: dangerous situations can be avoided, assistance can be given whenever and wherever needed, and the "cost of interaction" in terms of complexity of interaction can also be reduced.

Sustainable Industry (both sustainable 6G DT and 6G DT for sustainable industry): On the one hand, with massive twinning, the status of all components and participants in industrial processes can be precisely monitored, and jointly analysed. This helps people to thoroughly understand the unobservable patterns hidden beneath the complex processes, e.g., the carbon trace and the energy consumption throughout the process chain. Therewith, the energy efficiency and sustainability of the industry can be improved. On the other hand, green communication and energy efficient computing solutions must be invoked to ensure the industrial deployment of 6G massive twinning itself being sustainable.

5.2 Novel HMI for mobile human-machine and human-CPE interaction

5.2.1 Concepts

6G is supposed to provide ubiquitous connectivity of ultra-high density with low latency. This is paving the road towards a massive twinning scenario, where numerous physical entities are not only connected with each other by the pervasive Internet of Everything, but also synchronized with their corresponding digital twins, respectively. The coupling between the physical world and the cyber world is becoming tighter and smoother than ever before.

However, there is still one vertex missing from the triangle that completes the 6G vision of Hexa-X: *the human world*. To complete this ultimate gestalt of 6G network that joins three worlds, a seamless docking between the human world and the physical world, as well as between the human world and the cyber world, must be introduced as the last piece of puzzle. Such a task demands solutions for human-

context-awareness, immersive mixed reality (MR), and accompanying digital twins, which cannot be sufficiently fulfilled by the conventional user interfaces mainly based on mouse-keyboard-screen (MKS) and audio devices. Innovative human-machine interfaces are therefore called for to empower mobile human-machine and human-CPE interaction in the next decade.

5.2.2 Applications and requirements

5.2.2.1 Applications

Latest technologies of biosensor, biometric data processing and multi-sensory feedback will enable industrial systems to percept, understand, and interact with human. Therewith, humans can be more deeply involved with the industrial processes by means of various emerging DT applications, which are including but not limited to:

Human-context-aware industry: capturing the physical and mental status of human beings, novel HMIs support the DTs of human workers in industrial environments with rich and comprehensive data, which can be collected, archived, and analysed by the DT of human. An interaction between the DT of physical environment and the DTs of human workers will further make it possible to estimate and predict the human behaviour pattern. Inappropriate operations and unexpected events can be therewith avoided to ensure high security.

Telepresence collaboration with multi-sensory mixed reality: the timely synchronization of status and context information from humans to their DTs have established a basis for real-time telepresence of multiple users over future 6G networks. Furthermore, with the latest HMI such as holographic vision and tactile feedback, a multi-sensory mixed reality service can be delivered to enhance the experience of interaction between the on-site people, the telepresent people, and the environment.

Accompanying DTs: with long-term archived and well brokered human information, a human DT will be able to provide real-time context-aware intelligent service to its corresponding person. Combined with novel multi-sensory HMI technologies, which not only ease the user operation of personal devices, but also enrich the perception of feedbacks from them, accompanying DTs that serve as virtual personal assistants are passing from sci-fiction to reality.

5.2.2.2 Requirements

Safety and health-friendliness: as a human-centric system, 6G takes the safety and health of humans as its core value and priority. Axiomatically, any technology with potential risk of damaging the health of users by itself shall never be considered before being eventually verified as harmless. In addition to that, due to the particularity of industrial environments, every candidate HMI solution shall be evaluated in the context of the specific use scenario, to make sure that it does not significantly distract or restrict the user from the main task.

Comfort and convenience: beyond the basic requirements of health and safety, user's comfort and convenience shall be considered. This generally requires a compactness of the essential hardware and rejects most (if not all) invasive interfaces. For example, to recognize facial expressions and gaits, graphic solutions based on cameras [KWL+09, CSG+03] are preferred in this sense to solutions based on electromyogram (EMG) signals that can only be measured using electrodes [LBP09, GS14]; should the EMG signals be measured, surface electrodes are much preferred than needle electrodes.

Dependability: for applications intolerant to outages, 6G intends to deliver guarantees for multiple E2E capabilities, such as achievable data rate, maximum E2E service latency, bounded jitter, and E2E packet reliability with robust mobility. This is known as the concept of dependability [UER+21], which is especially critical for use cases such as human-machine interaction and automation. As an essential stage of the E2E service chain in 6G-driven industrial DTs, the HMI must cope with these requirements.

Electromagnetic Compatibility: massive twinning in industrial scenario inevitably leads to a complex electromagnetic environment, where strong interference and noise may probably occur, sometimes even randomly and without predictable pattern. The HMI solution therefore must be electromagnetically

robust to ensure its reliability. On the other hand, it shall not generate strong electromagnetic leakages that interfere with the wireless channels or other devices.

Sustainability: 6G takes sustainability also as one of its key values. This does not only mean that 6G shall foster improved sustainability in various societal domains, but also implies that 6G itself must be made sustainable. Unsustainable technologies, which depend on non-renewable resources, use pollutive materials, or generate high CO2 submission, shall not be adopted in 6G with no exception for the HMI.

Security and privacy: As 6G will, with its ubiquitous coverage and massive twinning, connect all things and people with each other and carry massive data that describe them comprehensively and in detail like never before, it also raises concerns in security and privacy to an unprecedented level. On the one hand, enhanced security measures must be taken to guard the user data unauthorized accesses and malicious operations by a non-trusted person or third party. On the other hand, the user data also need to be protected from possible inappropriate exploitation by the trusted ones, like the industrial verticals. For instance, the General Data Protection Regulation (GDPR) of the European Union prohibits data processing that can leak the user identity or lead to discrimination, with only a few exceptions under very strict rules. How to exploit human-specific data in DT systems (especially those with DTs of humans) while complying to such regulations must be taken into consideration when designing the HMI.

5.2.3 Key enabling technologies

5.2.3.1 Sensing the human mental status

Industrial processes can be greatly impacted by the mental status of the human participators, not only regarding the quality of output/product, but more importantly, regarding the safety of systems and humans themselves. The most critical mental signals to be measured are listed below.

Comprehension: in industrial scenarios, especially regarding human-machine collaboration, it is usually important for the intelligent systems to confirm that the human has perceived and understood the critical information they provide, such as a notification, an inquiry, an instruction, or a warning.

Concentration: distraction from the task in processing in industrial environment can easily lead to operation failures, which may cause serious losses and dangers.

Fatigue: fatigue generally reduces the level of concentration. It also weakens human's strength, agility, precision, and endurance in physical tasks, degrading the human cognitive abilities in all aspects. In addition, continuous working under fatigue is likely to damage health.

Emotion: negative emotions such as depression, stress, anxiety and anger can regularly lead to distraction, and accelerates the increase of fatigue. With the significant impact on the human behaviour pattern, extreme emotions may also cause failures of human-related prediction algorithms.

Since direct sensing of the mental status is impossible, people have developed a handful of indirect approaches to estimate it from measurable physical features. Physical features that are typically taken as indicators include those discussed below.

Speech voice: emotions and fatigue can significantly change the tone, speed, volume and timbre of one's speech voice. Hence, they can be detected through vocal analysis. However, to continuously monitor the mental status of a person in this way, the person has to keep talking, which is not universally possible. Nevertheless, in some scenarios where the vocal signal is available, e.g., when vocal commands are used to control the intelligent system, voice-based mental status estimation can be well integrated.

Facial expressions: rich information about one's mental status can be mined from facial expressions and micro-expressions, which have been exploited by human beings for thousands of centuries as the most classical and reliable approach to emotion recognition. This approach requires images or videos

that capture most facial areas with a sufficiently high resolution and ideally in the front view of the person under analysis, which are easy to obtain in most industrial scenarios.

Galvanic skin response signal: shifting in the intensity of emotional states can cause effects in human's eccrine sweat gland activity, which change the conductance of skin. Such effects are almost immedia indicators of emotional arousal, and easy to measure with wearable devices.

Eye movements: eye movement has been used since long as a psychological signal for emotion recognition. It can be captured either from video by cameras and eye trackers, or from electrooculogram signals by specialized electrodes. Both can be integrated in wearable devices like VR/AR headsets.

Bioelectric signals: there are a variety of bioelectric signals widely used in psychological and clinical studies to identify human brain activities and mental status, common examples include electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), and electrocardiogram (ECG). They usually need to be sensed with specialized electrodes, which are not always convenient and comfortable to carry while working in industrial environments. Some examples of mental status sensing are listed in Table 13.

Table 13: Examples of estimating the human mental status over physical features measured over novel HMIs.

Mental signal	Physical feature	Reference
Comprehension	Facial expressions	[TWL+18]
	Galvanic skin response	[BB20]
	Eye movements	[LNL+16]
	Bioelectric signals	[RWS+16]
Concentration	Facial expressions	[SEK19, TYN+05]
	Bioelectric signals	[KIS19]
Fatigue	Speech voice	[GFW+06]
	Facial expressions	[KHX+18]
	Galvanic skin response	[YLB05]
	Eye movements	[SBK10]
	Bioelectric signals	[KTS+20]
Emotion	Speech voice	[HDM+14]
	Facial expressions	[LWC+17]
	Galvanic skin response	[LFZ+16, SPC+20]
	Eye movements	[SR17]
	Bioelectric signals	[SWA18]

5.2.3.2 Multi-sensory feedback:

Since vision and hearing are the most important sense of humans, visual and auditory user interfaces (UIs) will still indefinitely remain the dominating approach for machines sending information to humans. However, the traditional visual/auditory UIs based on text, two-dimensional graphics and audio cannot fulfil the requirements of future industrial DT applications like MR, immersive telepresence, and human-robot collaboration, and therefore must be extended.

Holographic vision: As the successor of text, image, video and 3D video, holography will be the fifth generation of visual user interface technology. Allowing such a visualization of DT, with no clear boundary sensed between the virtual objects and the real physical environment, holographic vision plays a key role in future DT applications including MR, immersive telepresence, and telecollaboration.

There have been already commercial holographic vision products released, which have also been applied in research works, such as the Microsoft HoloLens [UBG+20]. They have been proven sufficient as solutions to optically deliver MR. To achieve the target of teleinteraction and telecollaboration, however, there are still two main challenges to overcome. First, the closed-loop

latency must be minimized, considering the computing delay to update the holographic model of the DTs regarding human actions (e.g., pressing a button on the holographic projection of a machine), so as to realize a smooth teleinteraction experience. Second, when multiple users are involved, e.g., in telecollaboration scenarios, the holographic image/video must be highly synchronized for all users. Both challenges are related to the research areas of URLLC, information freshness, and time-sensitive networking.

Tactile: Alongside with vision, the mostly used human sense is that of touch, which is exploited in every physical contact with the environment or any object. Being partly processed by the spinal cord instead of the brain, touch is also the most agile sense of humans, responding significantly faster than the visual and auditory senses [NC12]. Since the concept of Tactile Internet being proposed in 2014 [Fet14], the telecommunication community has been struggling to achieve the target of ≤10ms E2E latency, which is the limit of tactile cognition for humans. In the wireless field, the URLLC use case was proposed to address this issue, breeding numerous research works over the past years.

Despite the significant progress in evolving the network infrastructure to support Tactile Internet, haptic solutions to generate agile and accurate tactile feedback are still far behind the requirements of future industry. The haptic interface is a critical enabler for industrial DT applications, not only because it enriches the user experience of sensing virtual objects with touch, but also as a necessary condition of reliable motion capture-based remote operation. Only with an accurate and fast haptic stimulus that vividly simulates the weight, resistance, and hardness of physical objects, it becomes possible for human users to appropriately exert and accomplish high-precision tasks such as tele-surgery.

Traditional solutions to provide tactile feedback are mostly extrinsic, i.e., they rely on instrumented environment with active devices, such as vibrators-embedded steering wheels for virtual automotive driving [DL15], or the vibration-based tactile simulation of a physical keyboard in [BCB07]. Such solutions are generally limited to local areas and specific use scenarios. To enable tactile feedback in immersive and ubiquitous teleinteraction with arbitrary objects, the intrinsic solutions alternatively try to augment the user rather than the environment and alter the user's tactile perception. This can be typically realized by wearable haptic devices, or through direct stimulations to the user's neurosensory mechanism [BP12]. Recently, there has been a new class of methods developed, using drones to flexibly generate tactile feedback [AKH+17, KKS+17].

Other senses: In addition, other senses such as smell and taste may also be useful in some industrial scenarios, but the physicalising of them remains a technical challenge [JDI+15].

5.3 Modelling the Impact of Human Presence on Industrial Deployment of Digital Twins

5.3.1 Motivation

Wireless communication with URLLC is paving the way for factory automation in Industry 4.0. The dominant traffic in these industrial environments is expected to be machine-to-machine traffic. Guaranteeing deterministic low-latency communication across the entire factory floor requires accurate modelling of rare failure events so that their impact during operation is understood and can be limited. In other words, we need to minimize the overall randomness in the system so that errors between models and measurements are minimal.

This is possible with future factories that are semi-controlled environments where accurate position information for mobile UEs such as AGVs is available. The exact location of fixed machinery is also known. We can also expect accurate prior knowledge of the UE characteristics and capabilities such as antenna beam patterns, and number of antennas and their locations and orientations on the device. With location information and knowledge of the propagation environment (shape of obstacles in the environment and their materials as well as the impact of these obstacles on radio propagation), one can therefore model the radio condition between the UEs and the network using a DT. Given a model of the factory floor and the locations of the UEs and the network nodes, the DT could use sophisticated ray-

tracing algorithms or electromagnetic solvers, or rely on a data-driven approach, or use a hybrid of both methods to predict the link conditions between the UEs and the network.

Without human presence, an industry 4.0 environment can be made rather predictable. For example, the UE mobility can often be planned fully in advance and modelled in DT (for example when the UE is attached to AGV or robot). The DT could be used to evaluate the radio conditions between the network and the UE and proactively plan UE locations to avoid outages.

5.3.2 Modelling humans in a radio-aware DT

However, if humans are in the environment, their behaviour is much more unpredictable. Compared to the movements of a UE mounted on an AGV, whose location is known or can be controlled, precise location information of a UE that is being used by a human (e.g., handheld device) is harder to obtain. In addition, the movement of such UEs are inherently random and cannot be controlled. This makes it difficult to guarantee low-latency communication across the entire factory floor.

Humans will also affect radio communication even when they are not active users (i.e., not using any UE, e.g., handheld) since they can block the radio links effectively, especially when the radio network is using very narrow beams. The propagation effects with humans, especially at mmWave and THz frequencies, also depend on the type of clothing worn, the position of the limbs, and the shape of the clothing on the body. Consequently, an exact deterministic model for the effect of humans on the propagation environment is not possible since human (body and body parts) movement is mostly unpredictable. However, the limbs can be assumed to be limited to a certain area around the body, and the shape of the clothing on the body is very difficult to be deterministically modelled.

We must therefore rely on statistical models, where characteristics such as clothing are abstracted out [MAF19]. However, some predictions for the parameters of the statistical models are possible, e.g., given the current position, humans are likely to be in a limited area around that position. Similarly, given the current position of limbs, a likelihood for limb positions around that position can also be modelled. Note that the current limb position and human body location can be obtained through on-body sensors or cameras placed in the environment, and would qualify to be a novel human-machine interface for enabling a radio-aware DT (c.f. Section 5.2).

The accuracy of the statistical models can further be improved by personalizing the model parameters to the human on the floor. In other words, since we already know the employees working on a factory floor beforehand, the location of individual employees may be identified, and the statistical model parameters may be customized for their body types. However, to protect privacy, we require that such an identification is done with the individual's consent. The implementation may also be privacy-preserving if the information of the detected individual is discarded after being used for radio-resource management at the PHY/MAC layer and not retained anywhere else.

However, the Digital Twin can predict the effect of human presence and actions in an intelligent and pro-active way. Humans must be modelled as objects in DT and their position and possible actions (movements, gestures etc.) in certain geographical region at certain future time window have to be taken into account when planning how to use the radio resources in that region in that time window. This is critical with services having very high reliability targets. As an example, a simple hand waving can cause outage for service link with very high latency and reliability targets. This has to be taken into account when planning the beams and the frequency to be used. Also, enough redundancy and diversity can be prepared in advance if there is high probability for human blockage.

The statistical model of the human can be used to help proactively adapt to the varying link conditions and prevent outages. We can utilize the statistical model of the human to estimate reliability (from the DT) for a pair of beams selected at the transmitter and receiver at a given human body location and orientation. The reliability estimate can then be used to proactively adapt to varying link conditions via aggressive link adaptation and beam-switching to avoid blockage and/or improve channel conditions.

In scenarios that demand high reliability, a robust approach can be used to perform beam-alignment/switching with the statistical model of a human on the factory floor. We first obtain a set of

candidate locations and orientations of the human for the next instant given the current location and orientation of the human and his/her limbs. We also generate a set of candidate beams at the transmitter and receiver. Then, for each candidate beam-pair and each future location/orientation, we evaluate an estimate of the packet error (or any other relevant KPI) with the DT. We drop beams from the candidate set that does not meet the KPI. Now that we have a smaller candidate set, we optionally perform measurements over this smaller beam set and then pick a beam that satisfies the KPI.

Current work is ongoing in generating a statistical model of the human body and then using this model for predictive radio resource management.

5.4 Massive Twinning with Human in the Loop

Massive twinning is envisioned as one key use case for 6G, involving the massive use of digital twins for representing and controlling actions in the physical world [Hexa-X D1.2]. The use case relies on a real-time and accurate/exact virtual representation of the physical and human world where the role of humans in the loop is crucial. The importance of human-in-the-loop is controllability, decision-making, operations adaptation and corrective maintenance [YEZ+20] [YKJ+21].

Massive twinning will be deployed in many ways in industrial environments. One way is the utilization of real time data from multiple sensors, production processes and robots/AGVs by digital twins for detecting and mitigating anomalies in real systems. The information provided by these digital twins can help to perform the necessary mitigation steps in case of anomalies and increase efficiency through reconfigurations of the communication system and dynamic adaptations of the production process. This case combined with collaborative robots use case from [Hexa-X D7.1] will be the main focus of this section.

Digital twin systems must adapt to any new conditions and situations occur in real world, leveraging on extremely reliable low latency and high throughput networks, in order to succeed real-time digital/virtual representation. However, keeping the human in the loop for control of, and decision making on, robotic systems is still essential for various reasons such as complexity of systems or even legal constraints. Human-robot interactions are mostly limited to the labs and expert operators, with difficult interfaces that make them unusable for most non-skilled users. Humans should interact with the system in an intuitive and easy way, without being overwhelmed with unnecessary data and actions. Thus, appropriate user-centred designs have to be analysed and tested for human in the loop interactions. These patterns and techniques for human in the loop are being defined, designed and prototypes are being created and validated.

Initially, it is defined how the system cooperates with the human, what actions should be taken and what challenges and problems will be solved by the human. This design is validated every time the graphical user interface is used to influence the final design of an intuitive human-machine system by analysing the perception of interaction and ease of use. In addition, Artificial Intelligence and ML tools (i.e., diagnostic and functionality allocation components described in Section 4.2) gather information and provide alerts, suggestions and actions that the user can see and take control of whenever is appropriate. In particular, we studied the massive twining of three robots cooperating and conducting some industrial tasks (see Section 4.2.2). Each robot (LoCoBots) has more than 11 servo motors, 5 sensors, battery and computer with other internal components (CPU, RAM, storage, antennas, fans, etc.). The information of all these elements is collected, analysed, including indications of failures or sub-optimal operation and used by the digital twin (some of them are available for the end user as well). By introducing multiple robots and sensors and processes from the industrial environment, we created massive twinning. In addition, it is introduced an HMI via desktop or VR application developed for this case, to monitor, supervise, teleoperate the robots and solve any problems that may occur. Specifically, a user is able to start, stop, control and move the robots and robotic arms, view images or videos from robot onboard cameras and static cameras inside the environment and so on. Figure 5.2 and Figure 5.3 illustrate the data provided by the digital twin and the way, either through keyboard/mouse/touch or even with specialized VR glasses can view/control the industrial environment.

Some of the main challenges of human-machine interaction are:

- Deal with changes in the role of humans, from planning, configuring, executing, and monitoring robotic tasks to a more automated and supervisory role.
- Introduce new human roles, from experts (i.e., developers, operators, system integrators, control system engineers) to non-experts.
- Avoid overwhelming human with data that don't need or understand.
- Avoid recurring to human's actions or confirmation for every task in robotic construction environments.
- Simplify the day-to-day use of HMI system; the design should be focused on three main interaction designs: supervision (the user can monitor and supervise the automated robotic tasks), notification (the system notifies the user of errors, exceptions, abnormal situations or any important information and events), and control (the user can take control of the system, to reconfigure the robots and networks or even solve a problem with the use of teleoperation).

At the base of this system's HMI design, there needs to be a configuration stage in which the human only setup the most important data in a simple and intuitive way. Finally, Augmented Reality (AR) and Virtual Reality (VR) to visualize robotic tasks should be explored for a better, simpler way to interact with the machines, adding a more immersive and integrative experience for the human. This will help humans to better observe the industrial environment with a virtual tour of the environment in which humans, robots and all other elements work and interact with each other to help remote users better supervise the industrial processes.

Features of the required HMI - digital twin interface are:

- Intuitive graphical user interface.
- 3D environment, robots and kinematic models.
- Real-time visualization of the industrial environment.
- Remote human-machine monitoring and control (human-in-the-loop interactions).
- Real-time notifications and feedback from the system and robots.
- Live video streaming from robots' onboard cameras or the whole industrial environment.
- Teleoperation of robots.
- Control of automated industrial robotics tasks.

Derived 6G KPIs from above features are the following (respective values are given in [Hexa-X D7.1]):

- Scalability for fulfilling latency requirements for massive number of devices (sensors, robots, AGVs, other industrial equipment).
- Latency for remote teleoperation and control of robots (1-50 ms [Hexa-X D7.1]).
- High speed for upload and download of robot generated maps and also enabling the use (or updates) of maps by robots in real-time without the need to have them onboard.
- Elasticity for on-demand dynamic resource allocation for digital twin operations in an elastic way (e.g., on-demand streaming of multiple robot cameras).
- Mobility support to seamlessly serve mobile end-devices (robotic services can run on cloud, edge and extreme-edge nodes).
- Robustness for seamless operation of digital twin enabled 6G application.

On the left side of Figure 5-2 it is shown the 3D digital twin interface with three robotic units having live streaming videos of cameras placed in the room as well as on each robot, showing the surrounding space and what each robot can see. On the right side of Figure 5-2 it is shown the alert that appears on the digital twin when there is a real time problem of a robot so that actions can be performed to overcome this issue. Finally, Figure 5-3 shows the VR interface through which the user can remotely visualize the industrial environment with the use of proper VR headsets.

Figure 5-2: Digital twin user interface and capabilities.

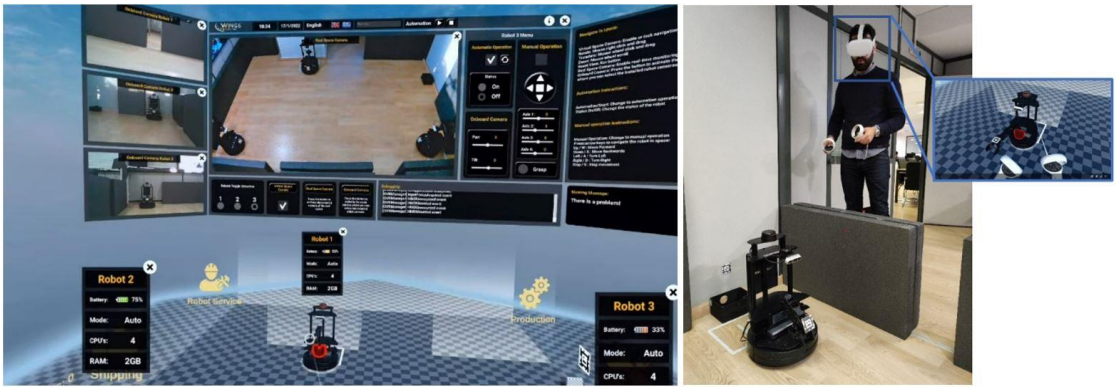


Figure 5-3: Digital twin with human interaction and VR capabilities.

5.5 Digital-Twin-Empowered Collaborative Robots

5.5.1 Introduction

Industry and world economy, in general, are being shaped by the great changes introduced by technology. Incredible innovations are impacting on all production processes, and it does not come as a surprise the fact that a concept like Industry 4.0 is now pervasive. In this process, that is recognized as the new industrial revolution, robots indeed play a central role. The authors of [GM19] state that three main directions of research can be identified in this context. One of these directions is that of industrial robotics, which has to do with the use of robots that can be re-programmed and exploited for more than one task [PL21]. Another direction is that of service robotics, which refers to the use of (possibly mobile) robots that can be autonomous, semi-autonomous or entirely tele-controlled and help humans in many different applications. The last direction is that of collaborative robotics, where machines known as “cobots” (i.e., collaborative robots, CRs), work by interacting closely with humans.

The research activity described herewith deals with the third direction, since it is the one that is considered as able to merge the advantages associated with the automatization of repetitive tasks (thanks to cobots) with the level of flexibility guaranteed by the involvement of humans on manual tasks [EMLU19], making the two agent categories work in a cooperative way. The research activity focuses, in particular, on a task that is of primary importance in the context of robotics, that is the programming one. Traditionally, industrial robots have been programmed offline, i.e., prior to robot operation: hence, resulting programs could not take human actions into account, leading to an overall poor flexibility of robot intervention. With CRs, though, programming can become an online, synchronous, operation, which can lead to more powerful, human-aware programs [EMLU19].

More specifically, several approaches have been developed to allow human operators change the robot’s programs: for instance, some authors experimented hand guidance [SN19], some others worked with voice and gesture commands [GSB+17] or with learning-by-demonstration paradigms [WHE+21]. It is rather common during the programming of a CR to record the path that its joints have to follow by

moving them physically using the hands rather than issuing commands to controls their position in space through, e.g., a controller device typically used with industrial manipulators [ASH21]. Nevertheless, there are situations in which manually guiding the CR to a specific position is not possible, e.g., because the position is not at reach of the operator's hands, it is dangerous, or it is not a viable option for some other reasons. Furthermore, creating a robotic program does not entail recording just robot's movements, but also other robot's operations (like opening or closing a gripper, etc.); in CR scenarios, an alternation between robot and human actions in a shared working space shall also be considered in programs being created. In such a scenario, emerging technologies such as eXtended Reality (XR) can help improving current methods for managing CRs. In particular, through XR it is possible to develop virtual environments where multiple users can, in principle, collaborate to carry out the programming task using 3D interaction methods and receiving visual feedback about their operations. For instance, with AR glasses, an operator co-located with the CR can see data regarding the working space or the robot's status. VR, in turn, could be used to program the CR in an intuitive way by means of 3D digital handles, possibly operated from distance. DT techniques could be exploited to reconstruct the real environment (robot's conditions, workpieces, operator's pose and actions, etc.) by leveraging both data provided by the robot, as well as information gathered by additional sensors (RGB, depth cameras and alike) located in the physical space.

Approaches to CR programming such as the one described in the latter paragraph, in which diverse categories of users (both co-located with the robot and remote) are involved and where multiple technologies (precisely, AR and VR) are leveraged, are considered to be extremely promising in the implementation of forthcoming collaborative robotics, as they could make the use of resources more effective and support developments in the perspective of so called "sustainable industry", e.g., by limiting the need for human operators to travel to the robot's site. Despite these advantages, it appears that the scenario illustrated above has not been studied yet in the context of real industrial settings. The reason might be found in the limitations that are currently exhibited by available network technologies, as dependability requirements regarding bandwidth and latencies that are needed to implement distributed XR applications could hardly be met today [FB21].

In order to address the limitations recalled above, an increasing number of works started to consider the opportunities that are being advertised by next-generation 6G networks [HS21, GPM+20, ZY21]. In particular, a feature that is considered of interest for collaborative robotics applications is the human-centric orientation of the 6G technology, which is expected to make humans gain a central role in all industrial scenarios encompassing robots [HS21].

Moving from this vision of 6G and from the envisaged use cases listed in [GPM+20], the present research activity aims to explore the creation of a programming methodology for CRs like the one previously described, by studying its requirements for what it concerns bandwidth occupation and end-to-end latency, and taking as a target reference the capabilities of upcoming mobile networks. More specifically, the focus is on a concept application scenario that strongly relies on XR technology and telepresence solutions, where two humans operating from remote sites need to cooperate to program a pick-and-place task on a CR. The robot needs to move a workpiece in a (e.g., dangerous) position which makes hand guidance not completely appropriate (Figure 5-4).

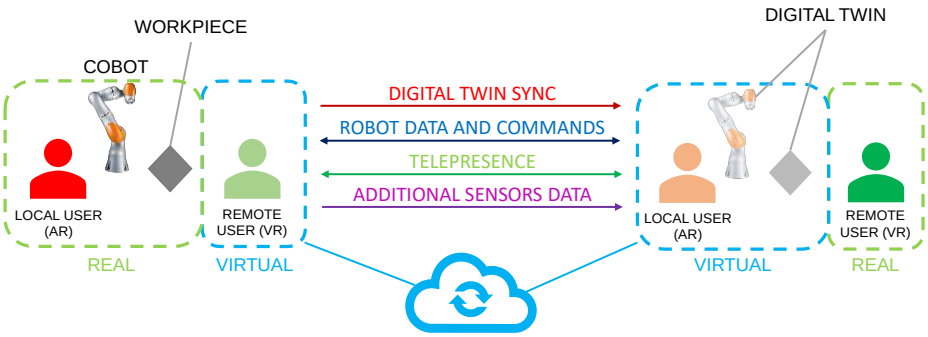


Figure 5-4: Concept scenario: Two users alternate to program a CR, one co-located with the robot and equipped with an AR HMD, the other one operating in a VR-based immersive environment.

One of the users is co-located with the robot (hence, it will be later referred to as "local user", or LU), whereas the other one operates from remote (hence, it will be named "remote user", or RU). The goal is to develop the pick-and-place program by alternating in the programming operation. The RU is immersed in a VR environment, in which he or she can interact with a DT of the robot and the workspace. The LU, in turn, wears an AR Head-Mounted Display (HMD) that allows him or her to see the physical space with digitally created contents overlaid. In this *alternate programming* approach, the two users can develop only the parts of the programs that can be better created in the environment they are operating into (the physical one, augmented with AR, or the digital one, in VR).

As it will be shown below, the results of experiments in which availability of a 6G network is simulated in a laboratory environment showed that the concept scenario could not be implemented using current 5G networks, but would become viable when considering next-generation mobile network technologies.

5.5.2 XR Telepresence-based Alternate Programming Approach

A system to experiment with the proposed concept of alternate programming is presented in this section.

5.5.2.1 Protocol

The alternate programming concept that has been analysed in this research is summarized in the steps described below, which refer to the interaction protocol depicted in Figure 5-5.

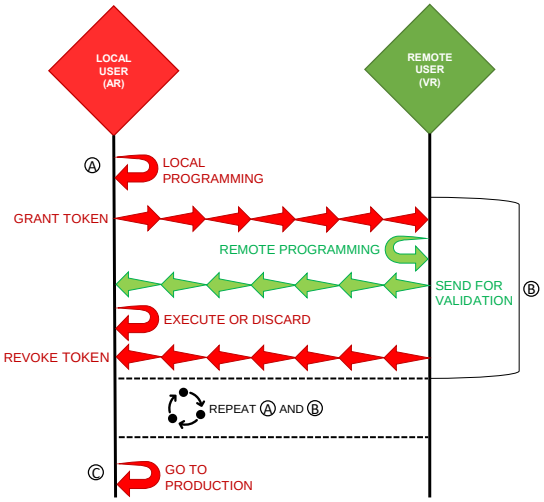


Figure 5-5: Interaction protocol considered for the collaborative programming approach, in which two distant users alternate on the program creation task using AR and VR tools.

The LU initiates the protocol by creating the first part of the program by interacting with the CR with typical means adopted in collaborative robotics (e.g., by guiding it with the hands, activating its tool using buttons available on the flange, etc.) as well as exploiting an AR-based interface. This interface lets the user record in the program the sequence of operations performed on/with the CR (movements of the end-effector and joints, timing, status of the tool, etc.), including also possible times at which an interaction between the robot and the agent is expected (e.g., one has to wait for the other to be ready, to provide an input, etc.) before moving on with the program execution.

At a certain point, the LU may decide, interacting with the AR interface, that it is time to pass the token for programming the robot to the RU, who has been simply a spectator in VR of the LU's actions till that point. Once acknowledged the receipt of the token, the RU can start controlling the CR (this situation is notified to the LU by a LED light on the robot's flange, if available, or in the AR interface). The RU starts its programming in VR by using the same functionalities that have been exploited by the LU in AR. In this case, however, the user sees the remote, physical working area in the virtual environment as a 3D reconstruction. More specifically, besides having the real CR recreated as a DT using data coming from its internal sensors, external sensors (like depth cameras) are additionally used

to realize a kind of telepresence embracing the two users. For instance, the RU can see a point cloud-based replica of the LU or of the workpieces present in the robot's surrounding. In order to limit the risks of damaging the CR or putting the LU in harmful situations, the RU is not allowed to have a direct control on the robot. That is, in his or her manipulations, the RU works on the digital replica of the robot. The advantage of working in VR is that the RU is not limited to work just within the physical boundaries of the working area, and can make the robot reach the intended places by moving virtually in the space and gaining optimal viewpoints on it (e.g., by using zoom in/out operations, moving the virtual camera that is used to frame the 3D space, etc.).

When that programming part is done, the RU can pass again the control to the LU. He or she checks that the program made by working on the DT is adequate (i.e., it does not include unsafe or unfeasible movements, it avoids collisions, etc.) by observing a preview, in AR, of a virtual robot that runs it. The LU can either discard the part of the program received by the RU, or accept and consolidate it in the overall program being created. The steps above are repeated with the LU and the RU collaboratively working on the program until the production process has been entirely coded and is ready to be executed. It is worth to point out that, while the program is being executed, the robot's motion cannot be fully controlled by the LU: he or she can only stop the CR intentionally using the safety stop or expect it to respond to a collision. In order to further contribute at smoothing the risks associated with these kinds of scenario, additional safety measures may be developed should information about the position of the LU and robot's status be available, e.g., implementing collision anticipation strategies.

5.5.2.2 Architecture

As said, this research activity aims to evaluate the feasibility of the proposed alternate programming concept from the perspective of mobile network capabilities. As 6G connectivity is not available yet, a prototype implementation was set up using 5GHz wireless and 100Mbps-to-10Gbps wired networking, realizing the architecture in Figure 5-6.

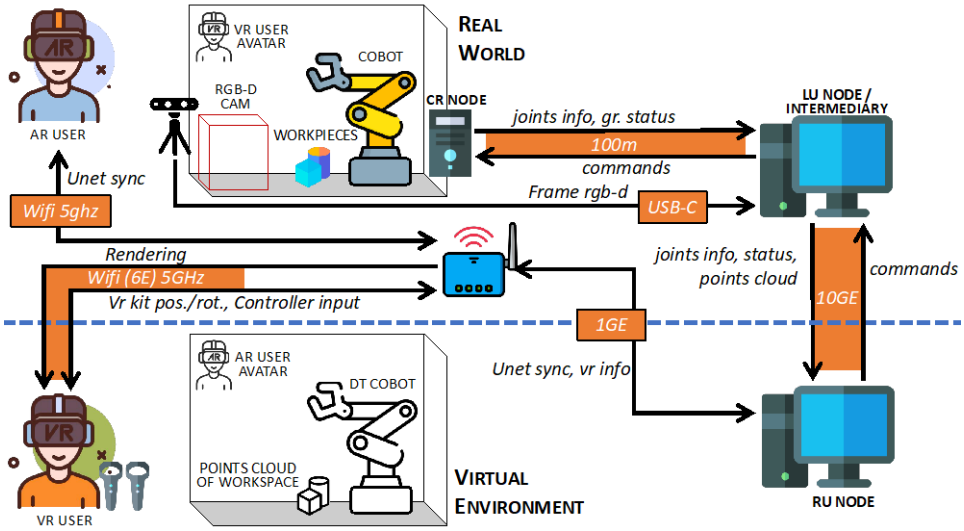


Figure 5-6: Architectural overview of the concept scenario that has been implemented for the analysis (prototype implementation set up using 5GHz wireless and 100Mbps-to-10Gbps wired networking).

The CR used in the experiments is a Kuka LBR Iiwa 14 R8201[7] (kindly provided by KUKA Roboter Italia Spa), operating in a working space of 4m × 4m × 3m. The robot's tool is an anthropomorphic hand by qB SoftHand Industry[8], mounted on the flange. An RGB-D camera is used to capture the

[7] KUKA LBR Iiwa: https://www.kuka.com/en-us/products/robotics-systems/industrial-robots/lbr-iiwa

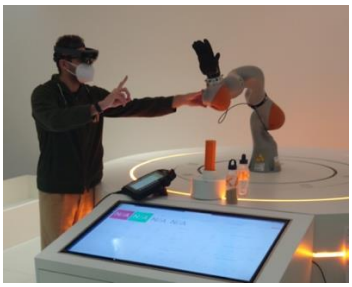
[8] qB SoftHand Industry: https://qbrobotics.com/products/qb-softhand-industry/

working space. The camera is connected on USB-C to a network node which is connected to the robot through the Ethernet X66 of the Kuka controller (later referred to as *CR node*).

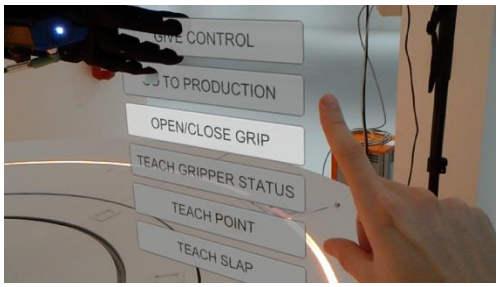
The architecture includes two computers that simulate the network nodes that are in charge of providing the users with necessary services. These nodes, later referred to also as *LU node* and *RU node*, are powered by Intel i7-8700 CPUs (6 cores @ 3.2GHz), Nvidia GTX 1080 GPUs and two-port 10Gbps Intel X540-AT2 network cards. In order to support the streaming of a RGB-D point cloud, a direct link between the computers was implemented. On the LU node there is also a so-called *Intermediary* broker that has been developed using the *ZeroMQ* library (later abbreviated ZMQ) and is set up as a proxy to tunnel the connection from the robot to the RU node and vice-versa. The alternate programming leverages a custom protocol that was created to define the messages to be exchanged for communicating information like the status of the CR, the issued commands, etc. The protocol relies on *protobuf* (v3.10) by Google for data serialization/deserialization. The RU node is in charge of simulation and rendering of the virtual environment. In order to emulate a thin-client paradigm that could be implemented on 6G in a remote rendering use case, the VR contents are visualized on an untethered HMD. The RU is thus allowed to move freely in the environment, as data transfer occurs on a 5GHz WiFi 6E network, through a router that is connected to the RU node. The latter node is also responsible for the synchronization of robot's DT and its status with the AR device on the LU side, which is connected to the same WiFi router. For the experiments, the LU and the RU operated from two different rooms at Kuka Roboter Italia Spa premises that were 25m apart. The dashed line in the figure illustrates this separation.

5.5.2.3 XR Platform

The XR platform, whose functioning is shown in Figure 5-7, has been implemented using Unity 2018.4. As it was designed as a multi-player application, the high-level network API provided by the chosen graphics engine (*UNET HLAPI*) was used to manage the networking aspects in a client-server fashion.

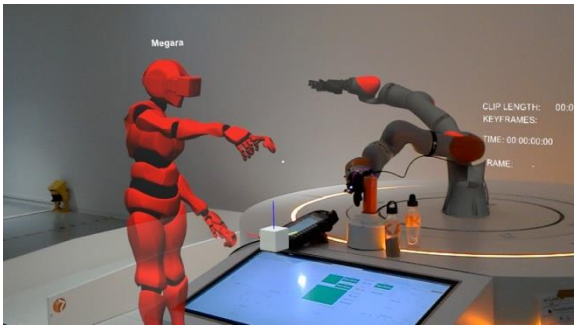


(a) LU with the AR HMD



(b) LU's viewpoint and AR UI

(c) RU using VR



(d) LU's viewpoint on the RU's replica

(e) RU's viewpoint and VR UI

Figure 5-7: Usages of the XR-based alternate programming approach.

The server could be hosted on a client or on a dedicated machine. In the experimental evaluation, it was deployed on the VR client hosted on the RU node. In order to manage the communication between the clients of the XR platform and the other elements of the devised architecture, *NetMQ* v4.0.1.6 (a native

C# porting of ZMQ) was exploited for opening the required channels connecting the UNET server and the Intermediary proxy component, as UNET is server-authoritative.

By means of the *UNET Network Transform* module, the position and orientation of all the clients was synchronized at a maximum frequency of 60Hz, enabling an interactive visualization of the other user. In order to smooth network updates, transformations were additionally interpolated. Synchronization was implemented using *UNET SyncVars*, whereas *UNET Hooks* were used to cope with events concerning SyncVar changes on the clients. Lastly, two additional UNET channels were established to enable voice communications between the clients with the *Dissonance Voice Chat* VOIP asset[9].

5.5.2.4 AR Application

The AR client, which was created as an application for the *HoloLens* HMD by Microsoft, leverages the *Mixed Reality Toolkit* (*MRTK*) for Unity to manage simultaneous localization and mapping, recognize gestures, and detect markers. Once the AR client connects to the server of the XR platform (the VR remote client), the user is asked to search for an ArUco marker (6 × 6, 10cm) that is placed on the base of the CR and is used to align the real and virtual elements; another marker is used to register the position and orientation of the RGB-D camera. As with UNET Network Transform it is only possibly to synchronize global transformations over the network, it was necessary to update the software to support also local transformations. Thus, the *Mixed Reality Playspace* could be aligned to the marker and then the relative coordinates of the AR HMD could be transferred with respect to it.

Once coordinate systems have been aligned, the AR user can see an UI in AR that displays information that comes from the CR (e.g., the rotation degrees of each joint). The user can interact with the robot either through hand-guidance, the controller, or the buttons on the flange, or directly via the UI, which includes widgets that allows him or her to open or close the tool, record a point in the path being programmed, pass the programming token to the other user or request it). As said when introducing the alternate programming paradigm, during the programming session the LU observes the operations of the RU, sees a preview of the programmed CR's trajectories and actions, and can accept or discard that program part before sending it to the robot for possible execution. In the following, the AR client running on the HoloLens will be named AR_{HMD}.

5.5.2.5 VR Application

The VR client was devised as an application that can be run on multiple devices, supposed that they are compatible with *OpenVR*. The *SteamVR* asset for Unity was used to implement common functionalities to manage, e.g., the camera, the controllers, the teleportation, etc. On top of that, the logic required for the collaborative programming was created. Once the hosting of the multi-player session has been started, the VR user automatically establishes a connection with the Intermediary broker. As soon as the client is connected, streams start to be received from the other platform components that are online, i.e., the robot and the RGB-D camera. The user cannot do anything else at this point besides requesting the token for programming the robot and wait for the AR user to connect. When the AR client connects to the UNET session, the VR can see the AR user, and vice-versa. More precisely, the AR user sees a full-body avatar representing the AR HMD, with one or two 3D spheres that indicate the positions of the user's hands (when detected). In the experimental setup, an *Oculus* (now *Meta*) *Quest 2* was selected as VR device. The VR rendering client that runs on it will be later referred to as the VR_{HMD}.

Once the AR client grants the programming token, the VR user is provided with an UI that allows him or her to control a digital replica of the CR, overlapped to the DT created from the status information received from the real robot. The digital replica is simultaneously displayed on the holographic glasses of the AR user. The VR UI, which is depicted in Figure 5-7 (e), is inspired by today's computer animation software suites, and allows the user to add or remove so called "keyframes" on a "timeline". When a sequence of robot operations (a part of a program) has been assembled, the VR user can "play"

[9] Dissonance Voice Chat: https://assetstore.unity.com/packages/tools/audio/dissonance-voice-chat-70078

them in a synchronous way on all the clients in order to let the AR user verify its appropriateness. If the latter does not spot any issues, the sequence will be sent to the robot and executed. Alternatively, the AR user may opt for discarding the sequence assembled by the other user. During sequence execution, the synchronized position of the LU is monitored to identify possible situations in which he or she gets dangerously close to the physical robot. If that occurs, a message is transmitted to the robot, which triggers a safety stop anticipating a potential collision. In both cases, i.e., sequence executed or discarded, the control is passed to the AR user. In the devised implementation, in order to simulate a remote rendering configuration, the Quest 2 was connected to the RU node that was running the VR client using *Air Link*. This functionality allows to use a standalone HMD in a tethered way without actually relying on a wired (USB-C) link. As said, in order to adhere to guidelines provided by Meta, the HMD was connected to a 5GHz Wi-Fi 6E router (ASUS RT-AX55), which in turn was connected using an Ethernet cable (1GE) to the RU node.

Besides seeing the DT of the CR and the digital replica of the AR user, the VR user is also presented with a real-time point cloud-based reconstruction of the remote working space collected by the RGB-D camera on the RU node and transmitted through the Intermediary broker. The position and orientation of the camera become available as soon as the AR user detects the associated ArUco marker. With this information, the VR client on the RU node can manage the incoming point cloud and visualize it locally. With this setup, information regarding the passive elements in the working space are collected and used for a remote reconstruction which could not be recreated otherwise.

5.5.2.6 CR Application

The CR application was coded using the Java language (v.1.7) for the *Kuka Sunrise OS* (v.1.16.2) leveraging the proprietary SDK components. Specifically, the *SmartServoMotion* (v1.16) and the *HandGuidance* (v1.0) components were exploited to control the robot's movement (when running a program or part of it) and to enable/disable the hand guidance teaching modality, respectively. To develop the networking layer, the ZMQ native Java implementation (*JeroMQ* v.0.4.3) was adopted. The CR node communicates with the LU node through the said messaging protocol, and manages it through a finite state machine (FSM) implementing the alternate programming concept. The anthropomorphic hand serving as a tool is attached to the robot's flange and configured using the *Kuka I/O Configurator* software available in *Kuka WorkVisual* (v5.0.5) suite; hence, it can be accessed by the XR platform and its status handled by the FSM.

5.5.2.7 RGB-D Camera Application

The RGB-D camera client was developed for a RealSense D435i5[10] device by Intel by resorting to the *librealsense2* (v2.49.0) SDK and its associated C# wrapper components. The application connects to the Intermediary broker on the RU node, accesses the camera over USB-C, obtains information required for the setup (intrinsic and extrinsic parameters of the sensors, frames per second, depth scale, etc.), then starts retrieving RGB-D frames and streaming them to the broker. Parameters are transmitted when the stream is opened and when a new client connects to it. The selected camera in equipped with a 1920 × 1080 pixels RGB sensor and a 1280 × 720 pixels depth sensor. The depth sensor can reach 90fps, whereas the RGB sensor is limited to 30fps. Each RGB channel is on 8 bits, whereas the depth channel is on 16 bits. The size of an RGB-D frame as well as of a second of stream are given in Table 14.

Table 14: Intel D435i data stream: Size of the RGB-D channels, for one frame and for one second of data (1920 × 1080 pixels for RGB, 1280 × 720 pixels for depth, 30fps).

	R	G	B	D	Total
1 frame	15.82Mb	15.82Mb	15.82Mb	14.06Mb	61.52Mb
1 second	474.60Mb	474.60Mb	474.60Mb	421.87Mb	1.80Gb

[10] Intel RealSense: https://www.intelrealsense.com/depth-camera-d435i/

According to Table 14, streaming one second of raw RGB-D data requires a rough bandwidth of 2Gbps. Even though the uplink limit of 5G networks that are under deployment is theoretically of 10Gbps [5GObs21], real performance won't probably reach that value [ROOT20]. Hence, 6G could be key to the deployment of such an architecture, thanks to its expected 1Tbps maximum bandwidth.

To physically simulate a 6G network, a wired 10 Gbps network was used for connecting the RGB-D camera client, the Intermediary broker component, and the VR client. Specifically, the broker and the camera client were run on the LU node, which was linked to the RU node on the VR client using a 30m-long CAT-6 Ethernet cable.

Once a depth message is received, the XR platform relies on an external module coded in C++ for creating the point cloud and its associated UV map, by implementing the algorithm that is openly included in the librealsense2 SDK. For speeding up the process, the process is multi-threaded, splitting the depth frame into eight subframes and handling them in parallel.

5.5.2.8 Deployment on 6G

The architecture in Figure 5-6 could be reworked as illustrated in Figure 5-8 to account for the features of a 6G network. As it can be seen, the components described in the previous sections are now directly connected to the 6G network using ad hoc adapters. Furthermore, the characteristics of the network allow for the deployment of services devoted, e.g., to remote rendering and to safety monitoring, on remote machines.

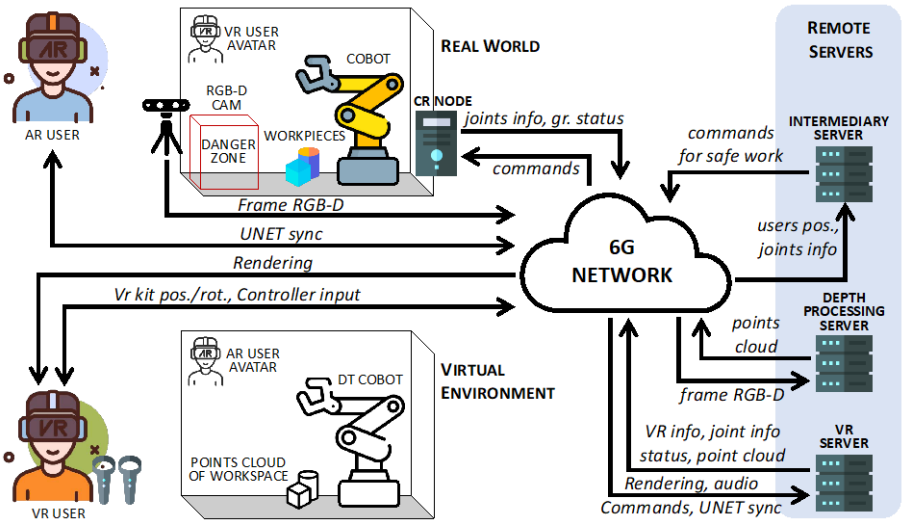


Figure 5-8: Conceptual architecture of the XR platform on 6G.

5.5.3 Experimental Results

In the following, the results gathered from the execution of an alternate programming task selected as a use case for the experimental evaluation of the devised approach in the laboratory setup introduced in the previous section are presented. In order to reproduce possible, real-world operating conditions and stress the platform under such conditions, two users were asked to repeat the task 10 times working at least 15 minutes on the programming in each repetition. The experiments were carried out at KUKA Roboter Italia Spa headquarters. A video is also available [Pol22].

The devised task encompassed a fictional pick-and-place production process. In order to simulate the presence of a region in the working space that cannot be reached by a LU co-located with the robot using hand-guidance, the workpiece (a bottle, in the specific case) was positioned on a 3D-manufactured support mimicking, e.g., a hazard zone in the operating range of the robot. The users' goal was to create this program: 1) the robot is initially configured with the tool (the hand) in safe position; 2) the tool is opened; 3) the tool is moved to pick the workpiece; 4) the tool is closed (to grab the workpiece); 5) the workpiece is moved outside the hazard zone where it can be reached by the LU; 6) the workpiece is

released in mid-air (place operation) as soon as the robot receives an input (a slap on the flange) by the RU; 7) the LU gives the input again and the cycle can be repeated from step 1.

For each program execution, statistics about bandwidth and latency between the network nodes were collected using custom probes in the developed software and common monitoring tools.

As for the VR rendering, the average bandwidth measured between the VR_{HMD} and the RU nodes was around 72Mbps, with a latency between a movement of the HMD and the updating of the visualization of $L_{VR\ HMD-RU}$ = 57ms (min 55ms, max 69ms). This result is in-line with the parameters configured for the Air Link and the reference performance advertised by Meta. As for the UNET layer, the latency between the visualization of AR contents and the VR UNET synchronization on the RU node was around $L_{AR\ HMD-RU}$ = 31.4ms (min 23.5ms, max 84ms); bandwidth requirements were rather low, on average equal to $B_{AR\ HMD-RU}$ = 100.8Kbps (min 1.9Kbps, max 228.8Kbps).

It has to be noticed that for obtaining the robot's DT visualization latency as perceived by the LU in AR, the delay above should be summed up to the delay required for the messages containing the CR's status information to arrive at the RU node. In the experimental evaluation, this latency was found to be equal, on average, to L_{CR-RU} = 66.7ms (min 23.3ms, max 148.3ms). A comparable latency is associated also to the receipt of a command issued in the AR interface on the CR node side.

5.5.4 Discussion

Below, results obtained with the experiments presented in the previous section will be discussed by considering network requirements set by the investigated alternate programming paradigm.

5.5.4.1 Bandwidth Constraints

Bandwidth requirements varied significantly for the connections used in the devised architecture. As for the LU–RU connection, that is managed using the UNET library, occupation was very low (~ 100Kbps) but also quite fluctuating over time, mainly because of the VOIP channel and the messages between clients and server for the synchronization of the XR devices. Regarding the CR–RU connection, the bandwidth occupation was even lower, as the connection was only used to transmit the messages about robot's status and, rarely, for commands issued and for program parts sent to the robot. With respect to the connection between the RGB-D camera client on the LU node and the RU, bandwidth occupation reached peaks of 2GBbps, a value that corresponds to the size of the stream to be delivered to the VR client on the RU node for recreating the point cloud of the working space.

Based on obtained results, it appears that mobile networks presently available would not support the required bit rates. In fact, 4G networks can cope with bandwidth occupations up to 150Mbps [Digi21], whereas today's 5G networks peak bandwidth can reach 1Gbps only in rare cases [ROOT20] (far below the performance of 10Gbps that would be theoretical possible [5GObs21]).

In such a scenario, 6G can be considered as a revolutionary technology for the implementation of use cases like the considered one, as given the theoretical upper bound of 1Tbps on bandwidth [Keys21] it would be possible to reconstruct entire physical spaces required for high-quality, interactive DT and telepresence applications. This peak performance would even make it possible to use more than one RGB-D camera at the same time, without impacting on end-to-end latency with possible compression and decompression operations required on lower bandwidth channels.

5.5.4.2 Latency Constraints

Results regarding latency showed, like in the case of bandwidth, quite an important variability. The measured latency between the AR_{HMD} node and the RU node (that is related to the UNET layer) was around 30ms; this value is mainly since, in the experiments, a WiFi connection was used, as this is the only network interface that is provided by the selected AR HMD. It is worth noticing that, to get the overall latency between the AR_{HMD} node and the VR_{HMD} node, the $L_{VR\ HMD-RU}$ latency of approximately 60ms due to remote rendering should be added too.

Regarding the latency between the CR node and the RU node, obtained results were numerically worse: even though a wired Ethernet connection was used, latencies between 23ms and 150ms were observed. In-depth investigation suggested that these counter-intuitive results were not due to the network configuration being considered, but to limitations of the Java Sunrise OS used on the robot. To mitigate this issue, alternative ways to communicate with the robot should be developed. Because of the above performance caps, the latency between the AR_{HMD} node and the CR node (corresponding to the sum of the above delays) was found to be around 100ms, on average. Given the fact that robot's movements can be relatively slow in the considered programming scenario, the overall network performance allowed anyway the proper functioning of the devised latency-dependent safety stop.

As for the bandwidth, replacing the WiFi connection with a mobile network channel could allow to reduce most of the delays above, especially between the LU node and the RU node. In particular, although the use of available 4G networks would worse performance given typical latency values around 50ms [Digi21], with 5G networks an average latency around 10ms could be already obtained [Keys21]. With 6G, delays between the AR_{HMD} node and the CR node could be further reduced, enabling the implementation of safety-stop feature also with higher-speed robot's movements and leaving as a remaining critical issue the latency between the RU and the robot.

5.5.5 Concluding Remarks and Future Works

The research herewith described focused on presenting and analysing the feasibility of a new approach for the programming of cobots that could play a key role in the sustainable development of industry thanks to advancements envisaged for forthcoming mobile communication technologies. In the devised interaction paradigm, a LU equipped with an AR HMD and co-located with the CR, on the one side, and a RU endowed with a VR HMD and seeing a 3D reconstruction of the physical space in with the LU in working into, on the other side, alternate in the creation of the robot's program. The sensors available on the robot are exploited, together with a depth camera, to implement a DT of the working space as well as of the workpieces and the LU, which allows the RU to safely operate on a VR-based representation on the real world.

The proposed paradigm sets important requirements on the amount of data to be exchanged and on the delays for data receipt, which may clash with the performance of mobile networks that are available today. To deal with such constrains, this research activity considered the advantages that are expected to be brought by the deployment of 6G technology. With the aim of quantifying the requirements set on the network by the devised XR telepresence-based alternate programming paradigm in a realistic setup, a laboratory implementation has been developed using network technologies that are currently available. The results regarding bandwidth and latency indicate that realizing the proposed paradigm with current 5G networks is not totally feasible (especially because of the huge amount of data to be transferred), proving though the theoretical possibility to implement it using 6G technology.

Based on research carried out so far, envisaged work for the future could encompass widening the experimental analysis to other dimensions of the proposed paradigm like, for instance, the user experience. Furthermore, with 6G it could be possible to shift architectures like the one exploited in this prototype towards the thin-client paradigm, in which all the computation pertaining, e.g., the simulation or the rendering, are performed on nodes that generate the results to be visualized (e.g., onto AR and VR devices). To support this paradigm, additional investigations will be required to carefully determine the associated network requirements.

5.6 Network-Aware Digital Twins for Local Insight

5.6.1 Motivation and concept

In Industry 4.0 scenarios, utilization of digital twins focuses on capturing the industrial processes as well as the state of machines (or parts thereof) and goods that interact as part of these processes, referred to as an industrial digital twin. In case human interaction is part of this process, it is included in the respective model. The industrial digital twin is then utilized to optimize the respective factory, process,

or machine, e.g., by simulating different implementations of the process or by generation of control algorithms with AI. Traditionally, with wired networks, the network itself had limited impact on these aspects, as it was assumed to operate according to specifications. Moreover, in the past most machines and robots were often stationary. Factories for example, once planned, built, and operated did not change much over time. The consequence was a limited need for reconfiguration or optimization during operation.

With wireless communication and the additional envisioned capabilities of 6G, these assumptions no longer hold.

- On the one hand, the network itself can provide (part of) the execution environment for the digital twins or the applications/processes that are modelled within those twins, e.g., by providing Compute-as-a-Service capabilities or integration with edge compute capabilities such as MEC [ETSI-MEC]. Consequently, dynamic reconfiguration and optimization of the network and its services (resulting from the mechanisms proposed, e.g., in [Hexa-X D4.1] and [Hexa-X D6.1]) have an impact on the real-world performance and characteristics of the industrial process and need to be considered in the industrial digital twin.
- On the other hand, flexible mobile communication systems offer mobility for machines and robots within the factory, making a truly flexible Industry 4.0 scenario possible. With mobile machinery, AGVs, and robots it is important to know where all equipment is actually located, both for planning and efficient, secure, and safe operation as well as for monitoring of the environments. Next generation mobile networks will not just offer flexible communication but can also more accurately track the position of objects as well as sensing the environment, as outlined in [Hexa-X D3.1].

Therefore, respective exposure interfaces providing local insights into the network for industrial digital twins are required. One parameter for such an exposure framework is localisation and sensing information offered as a new capability in 6G systems. This information reflects the current setup and status of a certain environment and is one of the key information for a digital twin in an industrial scenario. Especially sensing features will pose totally new opportunities to monitor the surrounding and ongoings in a certain environment. From factories to hospitals or public events, understanding how humans, or machineries act will be the base for most decision processes. While security always plays an important role, in certain scenarios with heavy and dangerous machinery, safety assumes a prominent role (c.f. Section 4.5 on dependability attributes). Depending on the assessments and decisions made by operators (of factories, public train systems, etc.), safety measures must meet specific requirements based on regulations. Communication, localisation, and sensing must guarantee that no information is altered, lost or delayed and decision based on the digital twins' parameters can be trusted. In Hexa-X, this relation is being studied both, in WP7 and in WP3, Task 3.3, with a focus on the utilization of localization and sensing as a service.

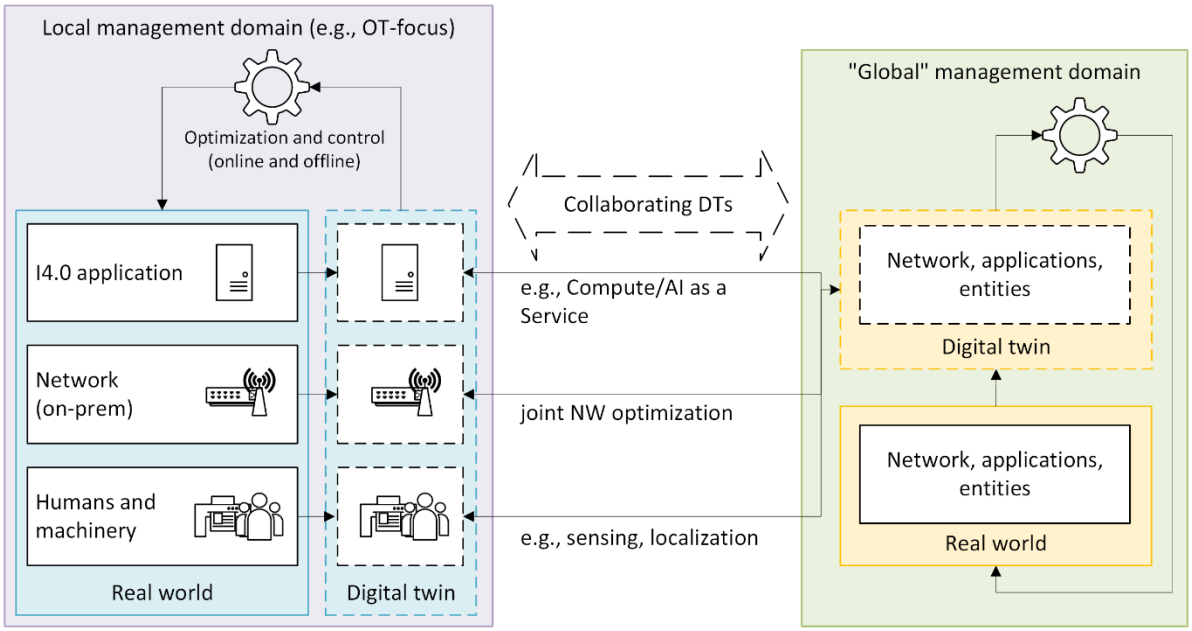


Figure 5-9: Concept of collaborating digital twins for local insight generation.

The concept is illustrated in Figure 5-9: a digital twin of a factory or a process, containing information about the respective I4.0 application and the involved machinery, is augmented with information on the network (on-premises) to form a local, network-aware digital twin (left part). This digital twin is then utilized for optimization and control of the real-world setup, in either an online or offline fashion. Here, "offline" refers to the possibility to alter the setup by deploying new infrastructure or by reconfiguring non-flexible parts of production lines if long-term benefits are expected. With the concept of collaborating digital twins, additional information is exchanged between the industrial digital twin and one or many digital twins of the 6G system (shown on the right) that provide additional local insights. One example is the utilization of localization or sensing functionality (with measurements being represented in the respective digital twin).

The exchange of information can occur bidirectionally: the industrial digital twin can, for example, provide information on local subnetworks managed in the local domain to the 6G digital twin(s) with the aim to further aid in (global) joint optimization of network resources or placement of functional components. One example is the optimization of trajectories of AGVs, as discussed in Section 3.2.

5.6.2 Experimental setup

For the realization of a network-aware digital twins for local insight, it is proposed to extend the dependability monitoring approach discussed in Section 4.5.4. In addition to time-series data being collected on network infrastructure elements and within applications, mobile sensors are deployed on AGVs and time/geo-series measurements for specific interactions between machines or between a machine and a human are collected. The application-level information from both fixed and mobile entities and the computing infrastructure being utilized for the industrial applications and for network functions of the communication system constitutes the local digital twin (left part of Figure 5-9). Network-level metrics are gathered on infrastructure or UE side (e.g., received signal strength indicators plus additional information being available at the radio unit, distributed unit, or centralized unit) and fed into a separate database (corresponding to the right part of Figure 5-9). Based on earlier works the measurement data are used to generate a radio map of the deployment. In a next step, this radio map is to be provided in a generalized form to the local digital twin containing the time/geo-series data for mobile entities, allowing the digital twin within the local management domain to optimize resource utilization (here, trajectories) without disclosing potentially business-confidential data to the network digital twin within the global management domain.

At the current stage, AGVs locate themselves using localization methods based on sensors like lidar, cameras, GPS, encoded wheels etc. [Hexa-X D3.1]. This requires translation into a point in a common coordinate system for the time/geo-series database, currently being realized by a web-based service running on the industrial edge environment. As outlined in [Hexa-X D1.2] it is expected that novel localization-as-a-service capabilities of the 6G system can be utilized by applications and services. The utilization of such a service for further extension of the local digital twin and the potential augmentation with sensor fusion from other sources (e.g., a separately deployed localization system on a factory floor) is under discussion with WP3 Task 3.3. Focus of the testbed and realization work conducted in this context is on the real-world implications and challenges for the outlined approach (e.g., considering the overhead introduced by the monitoring approaches, the uncertainties in localization estimates, etc.).

5.7 Digital Twins for Emergent Intelligence

5.7.1 Emergent Intelligence: Concept and Use Cases

As a concept originally proposed in a biological point of view [Hil88], emergent intelligence (EI) considers intelligence of animals, including humans, as an emergent behaviour, i.e., it is spontaneously originated from a large number of complexly interconnected and interacting simple units. For example, human intelligence originates from the brain that consists of massive number of interconnected simple neuron cells. Another typical example is the collection intelligence of gregarious insects such as ants and bees, which have only very simple behaving patterns with every individual, while exhibiting well developed intelligence as colonies.

The phenomenon of EI has also inspired engineers to develop bionic intelligent approaches, e.g., the particle swarm optimization method that iteratively improves a population of candidate solutions by moving them around in the search space, where the movement pattern of birds in a flock is imitated [KE95]. In the fields of complex systems and artificial intelligence, it has been studied to a certain degree. For instance, swarm algorithms have been developed over decades as a typical EI approach and effective metaheuristic optimization tool. Some researchers also study genetic programming and genetic algorithms from the EI perspective [Ang93, Ang94]. Further applications of EI have also been discussed in wide areas including logistics, resource allocation, job scheduling, privacy, among others. Nevertheless, EI has generally not become a very significant research field so far.

According to [Ang93], EI is distinguished from classical AI approaches mainly regarding the knowledge level and processing mechanism, as summarized in Table 15. Some descriptions to AI diverge from the modern point of view, e.g., decentralized AI technologies such as FL have been prospering over the past few years. Nevertheless, the fundamental difference between classical AI and EI still remains, i.e., AI takes a top-down logos that aims at an explicit target for the overall system, while EI is taking a bottom-up approach that relies on the interaction among independent agents to form implicit global pattern at the macroscopic level.

Table 15: Comparison between classical AI and EI.

Property	**AI**	**EI**
Task-specific knowledge	Integrated into problem solver	Separate from problem solver
Operators	Task-specific	Local and opportunistic
Processing	Centralized and global	Decentralized and representation specific
Knowledge representation	Explicit	Emergent
Knowledge content	Static	Dynamic and adaptive

Credit assignment	Task-specific and static	Representation specific, empirical, and dynamic

5.7.2 EI as a Novel Solution to Deliver Decentralized Intelligence

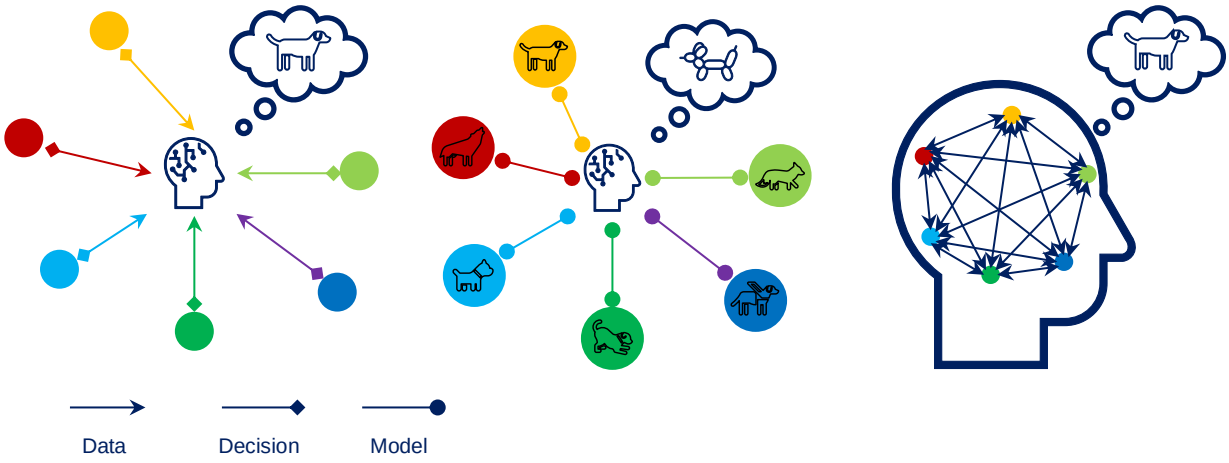


Figure 5-10: 2 Comparing the conventional AI solutions based on centralized Machine Learning (left) and Federated Learning (middle) to EI (right)

Compared to conventional centralized AI and FL solutions, EI exhibits the following features that benefit 6G applications:

- **Low computational complexity**
  Conventional AI solutions use highly complex abstract models, which can be typically represented by artificial neural networks, to fit massive data that describe the high-level statistical behaviour of the system regarding the global task. The process of model training is usually computationally expensive. FL partially distributes this training process to the agents, which locally specializes their own models based on the shared model provided by the central controller and send feedbacks to the controller to improve the shared model. Nevertheless, the overall computation effort is not reduced. In contrast, EI relies on no knowledge of the global task at any individual decision maker (though the knowledge is probably available at the system designer), but the interaction among massive agents that execute only very simple tasks. Therefore, it requires a significantly lower computational complexity than conventional AI solutions, which results in not only cheaper devices, but also a better sustainability.

- **Minimized latency**
  Without the demands of high computation effort or information aggregation, EI can significantly reduce the computation and communication delays in delivering intelligent services, which paves the way towards future intelligent industrial applications that require extremely low latency.

- **High robustness**
  EI provides a self-organized decentralized paradigm of decision making, which has no dependency on a central control unit. This grants EI with a special robustness against local malfunction at arbitrary agent, isolating the rest agents from the one(s) with malfunction so that they can still converge.

- **Privacy and security**
  Without aggregated storage and processing of user data at the central controller, EI only requires agents to exchange information with each other in awareness. This feature allows users

to only expose their data with trusted ones and reduces the risk of unwanted information leak. Thus, the data privacy is well treated. On the other hand, since the decision making is fundamentally distributed, without dependence on any shared model, it is impossible for attackers to globally bias user decisions by manipulating the model. This feature is implying a rich potential of EI application in DT-empowered CRs and collaborating DTs, which we introduced earlier in Sections 5.5 and 5.6, respectively.

- **Scalability**
  In conventional AI approaches, the central controller must globally collect data from every user, process the aggregated dataset, and distribute computation results to all users. This implies a linear increase in traffic volume and computing effort regarding the scale of networked system, which limits the system scalability. In contrast, the decision of each individual user in EI does not necessarily require the global information, satisfactory performance can be achieved with partial information exchanged among users within local subgroups. Indeed, a larger scale of the overall networked system usually improves the stability and convergence performance of EI.

5.7.3 6G Enables EI with Massive Twinning

In addition to the advantages provided by EI that 6G can benefit from, 6G can also support EI with the ubiquitous, massive, and reliable connectivity that it delivers.

As mentioned before, it generally requires a huge number of nodes participating in the networked system to realize EI, since the dimension and complexity of system are keys to any such emergent phenomenon. Envisaged to interconnect almost everyone and everything with its huge access capacity and full coverage over the entire planet, 6G will practically allow the networked system to arbitrarily upscale, which guarantees an effective emergence of intelligence.

Further than the capacity of user access, EI also requires reliable and timely communication to support efficient interaction between different agents. This will also be fulfilled by 6G with its ambitioned enhancement in all QoS perspectives regarding 5G.

Nevertheless, it shall be remarked that the requirements of system scale and communication efficiency can be usually opposite each other. When the number of agents increases within a limited coverage, to guarantee a full exchange of information between every two agents, the network access rate will geometrically increase therewith, which leads to a significantly raised collision rate or reduced radio resource per user, eventually resulting in higher latency or lower link reliability. Alternatively, the number of agents can also be increased by expanding the spatial dimension of the network. Even though the access density remains consistent in this case, the radio coverage can become an issue. Messages may have to be relayed, either by repeating agents or over the transport network, to be exchanged between two agents distancing from each other, which significantly increases the latency. Limiting the agent interaction range can be an effective solution, however, a degradation in convergence performance may be taken as the cost. Furthermore, in addition to the user plane data exploited by the agents to make decisions, a significant signalling overhead must be generated to setup and accomplish the communication sessions between agents, which can be of a very low energy efficiency that violates the sustainability expectation of 6G.

One promising solution to address this issue for EI is to deploy massive twinning. As the real-time status (raw knowledge), context information (abstracted/extracted knowledge) and semantic model (reasoning-level knowledge) of a physical object can be stored, analysed, and maintained at its digital twin, the decision engine of an EI agent can also be migrated from the physical device to its digital twin. Thus, the information exchange between different agents can be shifted from the physical radio environment to the cyber world. Every physical agent only needs to communicate with the cloud server where its DT is maintained, to upload its latest status and receive the decision made on the cloud. Thus, the massive radio signalling overhead can be mitigated, which significantly improves the radio resource efficiency and reduces the latency.

6 Conclusions

This report presented the progress, proposals, and intermediate results that have been conducted in the technical tasks within WP7 to address the need for special-purpose functionality in extreme environments. Initial solutions for *flexible resource allocation in challenging environments* (Task T7.2), *dependability in I4.0 environments* (Task T7.3) and *Digital Twins and novel HMIs* (Task T7.4) are presented, addressing the gaps identified in our first deliverable [Hexa-X D7.1] and following the work plan outlined there.

In this report, we provided a categorization of our initial solutions with respect to the targeted KPIs and KVIs, the architecture, and technical enablers based on the updated definitions and material in [Hexa-X D1.3]. We grouped contributions based on their cross-layer or cross-domain influences that are being utilized to achieve desired performance for specific use cases and specific extreme experiences, and highlight aspects that are shown in demonstrators. This mapping against KPIs and grouping into influence relations serves as a basis for future alignment with other technical work packages and will be updated with the final solutions of this work package in D7.3, due in May 2023.

Our initial solutions for *flexible resource allocation in challenging environments* include models for communication in factory environments, serving as foundation for interference management and link scheduling. These models consider the network-in-network nature of production lines in factories and the impact of mobility, both by controlled entities (e.g., AGVs) and humans. A second stream of contributions deals with the allocation and redistribution of resources for computation. A special focus is put on Federated Learning as one core building block for IoT use cases, especially involving resource constrained devices.

With respect to *dependability in I4.0 environments*, multiple approaches were shown that in some way converge communication and application instead of delivering ever-decreasing packet loss rates. By understanding more specifically the importance (*semantics*) of a packet, the network requirements can be greatly reduced, which enables to support an increased number of simultaneous real-time applications. In addition, practical approaches were demonstrated that highlight how dependability can be measured, quantified and guaranteed in industrial environments.

Regarding *Digital Twins and novel HMIs*, we presented the vision of an ecosystem of DTs for a human-centric realization of future industrial environments and discussed advances in HMIs that allow for a more seamless coupling between the physical and the digital world and a better capturing of biological aspects. Realizing parts of this concept, an implementation of collaborative robots utilizing a DT is discussed in detail. The potential of collaboration between different DTs as part of the overall ecosystem and foundation for trustworthiness is discussed, and the prospect of enabling Emergent Intelligence as a consequence of this collaboration is outlined.

All technical tasks in WP7 will continue the refinement of their contributions and report their final results in Deliverable D7.3, to be published in May 2023. This report will include results on the collaboration with other technical work packages on aspects related to security and privacy (WP1), utilization of positioning and sensing capabilities (WP3), AI/ML (WP4), architectural enablers for flexible topologies (WP5), and orchestration of resources (WP6), especially within the envisioned demonstrator. For these topics, dedicated workshops with the respective work packages and potential external stakeholders are planned in the coming months.

References

[38.801] 3GPP TR 38.801 V14.0.0 (2017-03) Study on new radio access technology: Radio access architecture and interfaces

[38.901] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," Release 16 (v16.1.0), 2019.

[5GObs21] 5G Performance European 5G Observatory, https://5gobservatory.eu/info-deployments/5g-performance/, accessed: 2021-12-01.

[AGM+10] Alizadeh, M., Greenberg, A., Maltz, D. A., Padhye, J., Patel, P., Prabhakar, B., Sridharan, M., "Data center tcp (dctcp)," In Proceedings of the ACM SIGCOMM 2010 Conference (pp. 63-74), August 2010.

[AJK21] A. Martínez-Alba, S. Janardhanan, and W. Kellerer "Enabling dynamically centralized RAN architectures in 5G and beyond." IEEE Transactions on Network and Service Management, 2021

[AKH+17] M. Abdullah, M. Kim, W. Hassan, Y. Kuroda, and S. Jeon, "HapticDrone: An encountered-type kinesthetic haptic interface with controllable force feedback: Initial example for 1D haptic feedback," in Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), pp. 115—117, 2017. doi: 10.1145/3131785.3131821

[Ang93] Angeline, Peter John. Evolutionary algorithms and emergent intelligence. Diss. The Ohio State University, 1993.

[Ang94] Angeline, Peter John. "Genetic programming and emergent intelligence." Advances in genetic programming 1 (1994): 75-98.

[ASH21] G. Ajaykumar, M. Stiber, C.-M. Huang, "Designing user-centric programming aids for kinesthetic teaching of collaborative robots," Robotics and Autonomous Systems 145 (2021) 103845. doi:10.1016/j.robot.2021.103845.

[AVK20] O. Ayan, M. Vilgelm and W. Kellerer, "Optimal scheduling for discounted age penalty minimization in multi-loop networked control," 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), 2020, pp. 1-7, doi: 10.1109/CCNC46108.2020.9045142.

[BB20] S. Bottos and B. Balasingam, "Tracking the progression of reading using eye-gaze point measurements and hidden markov models," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 10, pp. 7857–7868, Oct. 2020.

[BBA+21] G. Berardinelli, P. Baracca, R. O. Adeogun, et al., "Extreme Communication in 6G: Vision and Challenges for 'in-X' Subnetworks," in IEEE Open Journal of the Communications Society, vol. 2, pp. 2516-2535, 2021, doi: 10.1109/OJCOMS.2021.3121530.

[BCB07] S. Brewster, F. Chohan, and L. Brown, "Tactile feedback for mobile interactions," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07), pp. 159—162, 2007. doi: 10.1145/1240624.1240649.

[BDZ+19] Van Bemten, A., Đerić, N., Zerwas, J., Blenk, A., Schmid, S., and Kellerer, W., "Loko: Predictable latency in small networks," In Proceedings of the 15th International Conference on Emerging Networking Experiments And Technologies (pp. 355-369), 2019.

[BP12] O. Bau and I. Poupyrev, "REVEL: Tactile feedback technology for augmented reality," in ACM Transactions on Graphics, vol.31, issue 4, pp.1–11, 2012. doi: 10.1145/2185520.2185585.

[BT01] Le Boudec, Jean-Yves, and Patrick Thiran, eds. "Network calculus: a theory of deterministic queuing systems for the internet." Berlin, Heidelberg: Springer Berlin Heidelberg, 2001.

[BTM+20] Beutel, D.J., Topal, T., Mathur, A., Qiu, X., Parcollet, T., and Lane, N. D., "Flower: A friendly federated learning research framework," arXiv preprint arXiv:2007.14390, 2020

[CIFAR] A. Krizhevsky, V. Nair, and G. Hinton, "Cifar-10 (Canadian Institute for advanced research)" [Online]. Available: https://www.cs.toronto.edu/~kriz/cifar.html

[CSG+03] I. Cohen, N. Sebe, A. Garg, L. S. Chen, and T. S. Huang, "Facial expression recognition from video sequences: Temporal and static modeling," in Computer Vision and Image Understanding, vol. 91, issues 1–2, pp. 160–187, July–August 2003.

[Den12] L. Deng, "The mnist database of handwritten digit images for machine learning research," IEEE Signal Processing Magazine, vol. 29, no. 6, pp.141–142, 2012

[Digi21] Digitaltrends, M. Jansen and P. Beaton, "5G vs. 4G: How will the newest network improve on the last?," https://www.digitaltrends.com/mobile/5g-vs-4g/, accessed: 2021-12-01.

[DL15] L. Diwischek and J. Lisseman, "Tactile feedback for virtual automotive steering wheel switches. in Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '15), pp. 31—38, 2015.

[EGR+15] Emmerich, P., Gallenmüller, S., Raumer, D., Wohlfart, F., and Carle, G. "Moongen: A scriptable high-speed packet generator." In Proceedings of the 2015 Internet Measurement Conference (pp. 275-287), 2015.

[EMLU19] S. El Zaatari, M. Marei, W. Li, Z. Usman, Cobot programming for collaborative industrial tasks: An overview, Robotics and Autonomous Systems 116 (2019) 162–180. doi:0.1016/j. robot.2019.03.003.

[Eri21] Ericsson. "Radio Stripes: Re-thinking Mobile Networks". https://www.ericsson.com/en/blog/2019/2/radio-stripes, Accessed: 2021-07-29

[ETSI-MEC] ETSI GS MEC 002, "Multi-access Edge Computing (MEC); Phase 2: Use Cases and Requirements", V2.2.1, Jan. 2022.

[FB21] G. P. Fettweis, H. Boche, 6G: The personal tactile internet and open questions for information theory, IEEE BITS the Information Theory Magazine (2021) 1-1, doi:10.1109/ MBITS.2021.3118662.

[Fet14] G. P. Fettweis, "The tactile internet: Applications and challenges," in IEEE Vehicular Technology Magazine, vol. 9, no. 1, pp. 64-70, March 2014.

[FWP97] G. F. Franklin, M. L. Workman, and D. Powell, "Digital Control of Dynamic Systems," 3rd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1997.

[GFW+06] H. P. Greeley, E. Friets, J. P. Wilson, S. Raghavan, J. Picone, and J. Berg, "Detecting fatigue from voice using speech recognition," in 2006 IEEE International Symposium on Signal Processing and Information Technology, 2006, pp. 567–571.

[GIC+18] A. Garcia-Saavedra, G. Iosifidis, X. Costa-Perez, D. J. Leith, "Joint optimization of edge computing architectures and radio access networks." IEEE Journal on Selected Areas in Communications 36.11, pp. 2433-2443, 2018.

[GM19] R. Galin, R. Meshcheryakov, "Automation and robotics in the context of Industry 4.0: The shift to collaborative robots," in: IOP Conference Series: Materials Science and Engineering, Vol. 537, IOP Publishing, 2019.

[GPM+20] M. Giordani, M. Polese, M.Mezzavilla, S. Rangan, M. Zorzi, Toward 6G networks: Use cases and technologies, IEEE Communications Magazine 58 (3) (2020) 55–61.

[GS14] A. Gruebler and K. Suzuki, "Design of a wearable device for reading positive expressions from facial EMG signals," in IEEE Transactions on Affective Computing, vol. 5, no. 3, pp. 227–237, 1 July-Sept. 2014.

[GSB+17] P. Gustavsson, A. Syberfeldt, R. Brewster, L. Wang, Human-robot collaboration demonstrator combining speech recognition and haptic control, in: Proceedings of the 50th CIRP Conf. on Manufacturing Systems, Vol. 63, 2017, pp. 396–401.

[GSG+15] Grosvenor, M. P., Schwarzkopf, M., Gog, I., Watson, R. N., Moore, A. W., Hand, S., & Crowcroft, J., "Queues don't matter when you can JUMP them!". In 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15) (pp. 1-14), 2015.

[GTM+16] X. Ge, S. Tu, G. Mao, C. -X. Wang and T. Han, "5G Ultra-Dense Cellular Networks," in IEEE Wireless Communications, vol. 23, no. 1, pp. 72-79, February 2016, doi: 10.1109/MWC.2016.7422408.

[HDM+14] Z. Huang, M. Dong, Q. Mao, and Y. Zhan, "Speech emotion recognition using CNN," in Proceedings of the 22nd ACM international conference on Multimedia (MM '14), ACM, 2014, pp. 801–804.

[Hexa-X D1.2] Hexa-X, "Deliverable D1.2: Expanded 6G vision, use cases and societal values – including aspects of sustainability, security and spectrum", April 2021.

[Hexa-X D1.3] Hexa-X, "Deliverable D1.3: Targets and requirements for 6G - initial E2E architecture", March 2021.

[Hexa-X D2.1] Hexa-X, "Deliverable D2.1: Towards Tbps Communications in 6G: Use cases and Gap Analysis", June 2021.

[Hexa-X D2.2] Hexa-X, "Deliverable D2.2: Initial radio models and analysis towards ultra-high data rate links in 6G", Jan. 2022.

[Hexa-X D3.1] Hexa-X, "Deliverable D3.1: Localisation and sensing use cases and gap analysis", Jan. 2022.

[Hexa-X D4.1] Hexa-X, "Deliverable D4.1: AI-driven communication & computation co-design: Gap analysis and blueprint", August 2021.

[Hexa-X D5.1] Hexa-X, "Deliverable D5.1: Initial 6G Architectural Components and Enablers", December 2021.

[Hexa-X D6.1] Hexa-X, "Deliverable D6.1: Gaps, features and enablers for B5G/6G service management and orchestration", July 2021.

[Hexa-X D7.1] Hexa-X, "Deliverable D7.1: Gap analysis and technical work plan for special-purpose functionality", July 2021.

[Hil88] Hillis, W. Daniel. "Intelligence as an emergent behavior; or, the songs of eden." Daedalus (1988): 175-189.

[HR18] D. Harutyunyan, and R. Riggio, "Flex5G: Flexible functional split in 5G networks." IEEE Transactions on Network and Service Management 15.3, pp. 961-975, 2018.

[HS21] B. Han, H. D. Schotten, Multi-sensory HMI to enable digital twins with human-in-loop: A 6G vision of future industry (2021). arXiv:2111.10438.

[HSC19] HUBER-SUHNER, "5G functional split overview", December 2019. Available online: https://www.hubersuhner.com/en/documents-repository/technologies/pdf/fiber-optics-documents/5g-fundamentals-functional-split-overview

[HYJ+20] B. Han, S. Yuan, Z. Jiang, Y. Zhu and H. D. Schotten, "Robustness analysis of networked control systems with aging status," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2020, pp. 1360-1361, doi: 10.1109/INFOCOMWKSHPS50562.2020.9162929.

[IBM06] IBM "Autonomic Computing Whitepaper: An architectural blueprint for autonomic computing", Whitepaper, 2006. Available online: http://www-03.ibm.com/autonomic/pdfs/AC%20Blueprint%20White%20Paper%20V7.pdf

[IBN+19] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free Massive MIMO communications," in J Wireless Com Network 2019, 197 (2019). https://doi.org/10.1186/s13638-019-1507-0

[JDI+15] Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk, "Opportunities and challenges for data physicalization," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15), 2015, pp. 3227-–3236. doi: 10.1145/2702123.2702180.

[JSB+15] Jang, K., Sherry, J., Ballani, H., and Moncaster, T., "Silo: Predictable message latency in the cloud," In Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication (pp. 435-448), 2015.

[KE95] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95 - International Conference on Neural Networks, vol. 4, 1995, pp. 1942–1948.

[Keys21] Keysight, “6G: Going Beyond 100 Gbps to 1 Tbps,” Whitepaper, Available online: https://www.keysight.com/it/en/assets/7121-1152/white-papers/6G-Going-Beyond-100-Gbps-to-1-Tbps.pdf, accessed: 2021-12-01.

[KHX+18] S. A. Khan, S. Hussain, S. Xiaoming, and S. Yang, “An effective framework for driver fatigue recognition based on intelligent facial expressions analysis,” in IEEE Access, vol. 6, pp. 67459–67468, 2018.

[KIS19] H. Kaji, H. Iizuka, and M. Sugiyama, “ECG-based concentration recognition with multi-task regression,” in IEEE Transactions on Biomedical Engineering, vol. 66, no. 1, pp. 101–110, Jan. 2019.

[KKS+17] P. Knierim, T. Kosch, V. Schwind, M. Funk, F. Kiss, S. Schneegass, and N. Henze, “Tactile drones Providing immersive tactile feedback in virtual reality through quadcopters,” in Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA ’17), pp. 433—436, 2017. doi: 10.1145/3027063.3050426

[KTS+20] M. Kołodziej, P. Tarnowski, D. J. Sawicki, A. Majkowski, R. J. Rak, A. Bala, and A. Pluta, “Fatigue detection caused by office work with the use of EOG signal,” in IEEE Sensors Journal, vol. 20, no. 24, pp. 15213–15223, 15 Dec.15, 2020.

[KWL+09] W. Kusakunniran, Q. Wu, H. Li and J. Zhang, “Automatic gait recognition using weighted binary pattern on video,” in 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, 2009, pp. 49–54

[LBP09] D. T. H. Lai, R. K. Begg and M. Palaniswami, “Computational intelligence in gait research: A perspective on current applications and future challenges,” in IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 5, pp. 687–702, Sept. 2009.

[LBT+20] H. Ludwig, N. Baracaldo, G. Thomas, Y. Zhou, A. Anwar, S. Rajamoni,Y. Ong, J. Radhakrishnan, A. Verma, M. Sinnet al., “Ibm federated learning: an enterprise framework white paper v0.1,”arXiv preprint arXiv:2007.10987, 2020.

[LFZ+16] M. Liu, D. Fan, X. Zhang and X. Gong, “Human Emotion Recognition Based on Galvanic Skin Response Signal Feature Selection and SVM,” in 2016 International Conference on Smart City and Systems Engineering (ICSCSE), 2016, pp. 157–160.

[LML+19] Li, Y., Miao, R., Liu, H. H., Zhuang, Y., Feng, F., Tang, L. and Yu, M., “HPCC: High precision congestion control,” In Proceedings of the ACM Special Interest Group on Data Communication (pp. 44-58), 2019.

[LNL+16] J. Li, G. Ngai, H. V. Leong, and S. C. F. Chan, “Your eye tells how well you comprehend,” in 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC), 2016, pp. 503–508.

[LWC+17] Z. Liu, M. Wu, W. Cao, L. Chen, J. Xu, R. Zhang, M. Zhou, J. and Mao, J., “A facial expression emotion recognition based humanrobot interaction system,” in IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 4, pp. 668–676, 2017.

[LY18] L. Liu and W. Yu, "Massive Connectivity With Massive MIMO—Part I: Device Activity Detection and Channel Estimation," in IEEE Transactions on Signal Processing, vol. 66, no. 11, pp. 2933-2946, 1 June1, 2018, doi: 10.1109/TSP.2018.2818082.

[LY18a] L. Liu and W. Yu, "Massive Connectivity With Massive MIMO—Part II: Achievable Rate Characterization," in IEEE Transactions on Signal Processing, vol. 66, no. 11, pp. 2947-2959, 1 June1, 2018, doi: 10.1109/TSP.2018.2818070.

[MAF19] F. Mokhtari-Koushyar, M. Alamdar and M. Fakharzadeh, "Human Body Scattering Modeling and Measurement for Millimeter-Wave and 5G Bands," In Proc. IEEE Texas Symp. Wireless Microwave Circuits and Systems (WMCS), 2019, pp. 1-5.

[MBC22] M. Merluzzi, S. Bories, E. Calvanese Strinati, “Energy-Efficient Dynamic Edge Computing with Electromagnetic Field Exposure Constraints,” accepted at EuCNC & 6G Summit 2022, Grenoble, France

[NC12] A. W. Ng and A. H. Chan, "Finger response times to visual, auditory and tactile modality stimuli," in Proceedings of the international multiconference of engineers and computer scientists, vol. 2, pp.1449–1454, 2012.

[NCC17] Nguyen, T. D., Chiesa, M., Canini, M., "Decentralized consistent updates in SDN," In Proceedings of the Symposium on SDN Research (pp. 21-33), 2017.

[NO18] Nokia and Nokia Shanghai Bell, "Scenarios, Frequencies and New Field Measurement Results From Two Operation Factory Halls at 3.5 GHz for Various Antenna Configurations," document R1-1813177, 3GPP, Sophia Antipolis, France, Nov. 2018. [Online] Available: https://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_95/Docs/

[Ora21] Orange Press Release at MWC'21, 28th June 2021, available at: https://www.orange.com/sites/orangecom/files/documents/2021-06/MWC-2021-press-folder-en.pdf

[PBR+21] D. -T. Phan-Huy, D. Barthel, P. Ratajczak, R. Fara, M. d. Renzo and J. d. Rosny, "Ambient Backscatter Communications in Mobile Networks: Crowd-Detectable Zero-Energy-Devices," 2021 IEEE International Conference on RFID Technology and Applications (RFID-TA), 2021, pp. 81-84.

[PL21] F. G. Pratticò, F. Lamberti, Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at KUKA, Computers in Industry 129 (2021) 103446. doi:10.1016/j.compind.2021.103446.

[Pol22] Politecnico di Torino, Video, "Digital Twin Empowered Collaborative Robots", Available online: https://www.dropbox.com/sh/wepxflz9r6g3e8n/AAAou2bE3feLkG_1BmYw9V9Za?dl=0

[PPV10] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2307–2359, May 2010

[PSN+19] P. Popovski, Č. Stefanović, J. J. Nielsen, E. De Carvalho, M. Angjelichinoski, K. F. Trillingsgaard, and A. S. Bana, "Wireless Access in Ultra-Reliable Low-Latency Communication (URLLC)," in IEEE Transactions on Communications, vol. 67, no. 8, pp. 5783-5801, Aug. 2019, doi: 10.1109/TCOMM.2019.2914652.

[Root20] UK mobile performance in review 2H 2020, https://rootmetrics.com/en-GB/content/uk-mobile-performance-in-review-2H-2020, accessed: 2021-12-01

[RWS+16] Z. Rusák, N. van de Water, B. de Smit, I. Horváth, and W. F. Van Der Vegte, "Smart reading aid for detecting problems with reading fluency and comprehension," in Proceedings of the ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Charlotte, North Carolina, USA. August 21–24, 2016. V01BT02A039. ASME.

[SAM21] Samsung, "Virtualized RAN - Vol.2" Technical White Paper, April 2021

[SBK10] H. Singh, J. S. Bhatia and J. Kaur, "Eye tracking based driver fatigue monitoring and warning system," India International Conference on Power Electronics 2010 (IICPE2010), 2010, pp. 1–6.

[SCI20] Standardization Council Industrie 4.0, "German Standardization Roadmap Industrie 4.0"; March 2020. Available online: https://www.sci40.com/english/publications/

[SEK+19] P. Sharma, M. Esengönül, S. R. Khanal, T. T. Khanal, V. Filipe, M. J. C. S. Reis, "Student concentration evaluation index in an E-learning context using facial emotion analysis," in International Conference on Technology and Innovation in Learning, Teaching and Education 2018, 2019, pp. 529–538.

[SHC+10] J. P. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schöller, and P. Smith, "Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines," Computer networks, vol. 54, no. 8, pp. 1245–1265, 2010.

[SN19] M. Safeea, P. Neto, KUKA Sunrise Toolbox: Interfacing collaborative robots with MATLAB, IEEE Robotics Automation Magazine 26 (1) (2019) 91–96.

[SPC+20] I. Y. Susanto, T. Pan, C. Chen, M. Hu, and W. Cheng, "Emotion recognition from galvanic skin response signal based on deep hybrid neural networks," in Proceedings

of the 2020 International Conference on Multimedia Retrieval ICMR '20), ACM, 2020, pp. 341—345.

[SR17] R. S. Soundariya and R. Renuga, "Eye movement based emotion recognition using electrooculography," in 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), 2017, pp. 1–5.

[SWA+18] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, "ASCERTAIN: Emotion and personality recognition using ommercial sensors," in IEEE Transactions on Affective Computing, vol. 9, no. 2, pp. 147–160, 1 April-June 2018.

[TWL+18] C. Turan, Y. Wang, S. Lai, K. D. Neergaard and K. Lam, "Facial expressions of sentence comprehension," in 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), 2018, pp. 1–5.

[TYN+05] K. Takahashi, K. Yamada, T. Nakano and S. Yamamoto, "Method of detecting concentration on cellular phone call from facial expression change by image processing," in 2005 IEEE International Conference on Systems, Man and Cybernetics, 2005, pp. 3444–3448.

[UBG+20] D. Ungureanu, F. Bogo, S. Galliani, P. Sama, X. Duan, C. Meekhof, J. Stühmer, T. J. Cashman, B. Tekin, J. L. Schönberger, P. Olszta, M. Pollefeys, "HoloLens 2 research mode as a tool for computer vision research," arXiv:2008.11239, 2020.

[UER+21] M. A. Uusitalo, M. Ericson, B. Richerzhagen, E. U. Soykan, P. Rugeland, G. Fettweis, D. Sabella, G. Wikström, M. Boldi, M. H. Hamon, and H. D. Schotten, "Hexa-X The European 6G flagship project," in 2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), 2021, pp. 580–585.

[UVK+21] M. A. Uusitalo, H. Viswanathan, H. Kokkoniemi-Tarkkanen, A. Grudnitsky, M. Moisio, T. Härkönen, P. Yli-Paunu, S. Horsmanheimo, and D. Samardzija, "Ultra-Reliable and Low-Latency 5G Systems for Port Automation," in IEEE Communications Magazine, vol. 59, no. 8, pp. 114-120, August 2021, doi: 10.1109/MCOM.011.2001060

[WHE+21] Y. Wang, Y. Hu, S. El Zaatari, W. Li, Y. Zhou, Optimised learning from demonstrations for collaborative robots, Robotics and Computer-Integrated Manuf. 71 (2021) 102169.

[YEZ+20] Al-Yacoub, A., Eaton, W., Zimmer, M., Buerkle, A., Ariansyah, D., Erkoyuncu, J. A. and Lohse, N., "Investigating the Impact of Human in-the-Loop Digital Twin in an Industrial Maintenance Context" (2020). TESConf 2020 - 9th International Conference on Through-life Engineering Services, Available at SSRN: https://ssrn.com/abstract=3717797 or http://dx.doi.org/10.2139/ssrn.3717797

[YKJ+21] Yigitbas, E., Karakaya, K., Jovanovikj, I., Engels, G., "Enhancing Human-in-the-Loop Adaptive Systems through Digital Twins and VR Interfaces." 2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS) (2021): 30-40.

[YLB05] G. Yang, Y. Lin and P. Bhattacharya, "A driver fatigue recognition model using fusion of multiple features," in 2005 IEEE International Conference on Systems, Man and Cybernetics, 2005, pp. 1777–1784 vol. 2.

[ZEF+15] Zhu, Y., Eran, H., Firestone, D., Guo, C., Lipshteyn, M., Liron, Y., and Zhang, M., "Congestion control for large-scale RDMA deployments," ACM SIGCOMM Computer Communication Review, 45(4), 523-536, 2015.

[ZHK+21] Zhou, Z., He, M., Kellerer, W., Blenk, A., and Foerster, K. T., "P4Update: fast and locally verifiable consistent network updates in the P4 data plane," In Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies (pp. 175-190), 2021.

[ZY21] V. Ziegler, S. Yrjölä, "How to make 6G a general purpose technology: Prerequisites and value creation paradigm shift," in: Proceedings of the Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), 2021, pp. 586–591.