DETERMINISTIC6G

Use Cases and Architecture Principles

D1.1

The DETERMINISTIC6G project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no 1010965604.

Use Cases and Architecture Principles

Grant agreement number:	101096504
Project title:	Deterministic E2E communication with 6G
Project acronym:	DETERMINISTIC6G
Project website:	Deterministic6g.eu
Programme:	EU JU SNS Phase 1
Deliverable type:	Report
Deliverable reference number:	D11
Contributing workpackages:	WP1
Dissemination level:	Public
Due date:	M06
Actual submission date:	30-06-2023
Responsible organization:	B&R-AT
Editor(s):	Oliver Hoeftberger
Version number:	V1.0
Status:	Final

Short abstract:	This deliverable presents four visionary industrial use cases that highlight the need of deterministic E2E communication services in future 6G networks. In DETERMINISTIC6G these use cases are taken to drive the development of new architectural concepts, which will be the key enablers for realizing novel cyber-physical systems applications.

Keywords:	TSN, 6G, DetNet, Industry 5.0, Time Sensitive Communication, Extended Reality, Exoskeleton, Adaptive Manufacturing, Smart Farming, Key Value Indicators, Cyber-Physical Systems

Contributor(s):	Dhruvin Patel (EDD), Edgardo Montes de Oca (MI), Huu Nghia Nguyen (MI), Jose Costa-Requena (CMC), James Gross (KTH), Gourav Prateek Sharma (KTH), Leefke Grosjean (EAB), Joachim Sachs (EAB), Janos Harmatos (ETH), Oliver Höftberger (B&R), Franz Profelt (B&R), David Puffer (B&R), Drissa Houatra (OR), Francesca De Simio (IUVO), Giulia Bigoni (IUVO), Francesco Giovacchini (IUVO), Gargi Bag (ABB)

Revision History

V0.1	First draft
V0.2	Draft for final review
V1.0	Release version

Disclaimer

Executive summary

This report elaborates on visionary use cases that can be enabled by advanced deterministic end-to-end communication with 6G technology. The use cases highlight the trends in industry for ubiquitous integration of humans into industrial processes, new ways of collaboration and tele-working, increased mobility, and flexibility in production systems, as well as sustainability. An extended reality scenario and the support of human workers with exoskeletons are presented, which both benefit from the possibility of offloading computationally intensive tasks from the wearable device into the cloud. Multiple adaptive manufacturing scenarios illustrate the need for flexible configurations in automated factories, while functional safety must be maintained. Finally, smart farming is one representative use case of the mobile automation domain, which emphasises the progression towards highly autonomous mobile systems for domains like agriculture, mining, construction, embarkation, and many more. These use cases will act as guidance for the development of new architectural building blocks to enable better predictability, and in further succession, to achieve improved end-to-end determinism for converged wireless and wired network infrastructure.

Furthermore, the document introduces the concepts of key societal values and the corresponding key societal value indicators. These are intended to capture the gains and losses for the society and the need for sustainability, when using new 6G deterministic networking technologies. Based on the requirements of the use cases, their key performance indicators and key social value indicators, a deterministic service description will be defined during the project. An outline about the security analysis is given, which needs to identify new threats that may appear due to the introduction of new architectural concepts and to suggest possible countermeasures. For that purpose, a gap analysis with respect to existing standards and common security mechanisms is planned. Additionally, due to the impact of cloud computing on the deterministic E2E behaviour between applications, an elaboration of requirements for compute offloading is provided. Finally, a discussion about determinism in the scope of DETERMINISTIC6G is provided and the project-specific approach for enhancements to the 6G architecture is proposed. Implementing this approach will improve the predictability of the 6G network system behaviour, and thus, enables to perform proper actions that maintain the system within the specified system boundaries.

Contents

1 Introduction

The economic and social prosperity of our modern society is increasingly based on the technological advances that have been made possible by digitalization. Hardly any area of our daily life is unaffected by the trend to take advantage of the information and communication technology (ICT). It is expected that the digital transformation of life is even accelerating until the end of the current decade, as technical enablers mature and spark further innovations. This will result in new types of interactions between people, collaborations between humans and machines, and between autonomous machines. The perceived distance between people is reduced by immersive technologies like extended reality (XR), holographic communication, or the exchange of human-grade sensory information. Availability and usage of computational resources and data storage experience an evolution due to the introduction of cloud and edge computing. Digital twins can model entities and processes of the physical world, and based on that, cyber-physical systems (CPSs) can be created. In combination with machine learning (ML) and artificial intelligence (AI), novel strategies for analytics, optimization, and control are facilitated.

These trends are particularly relevant for the next generation of industrial production systems and lead to the fifth industrial revolution – also denoted as Industry 5.0 [LSW+22]. The importance of Europe's industrial production has also been realized by the European Commission, which added this sector to their strategic digital roadmaps [EU2020]. Industry 5.0 is fostering the inclusion of humans into industrial CPSs and the achievement of highest levels of system dependability, while at the same time, improving the sustainability of new technological developments. Some of these challenges can be addressed by the integration and enhancement of novel techniques for robot collaboration and tele-operation, by using new types of human-machine-interactions (HMI), for instance via Extended Reality (XR), or through the physical support of workers in their working processes (e.g., via smart exoskeletons).

One key enabler to bridge the digital and physical worlds in mobile and dynamic scenarios are smart 6G networks, closing the gap to other ICT services. At the same time, upcoming 6G networks will have to satisfy the growing societal need for sustainable solutions. To better understand the impact and value of 6G solutions for the society, additional research needs to be done. The definition of Key Value Indicators (KVIs) shall create new possibilities to amplify positive social and environmental effects and to mitigate losses in social values.

1.1 Goal of the DETERMINSITIC6G project

DETERMINISTIC6G has the goal to pave the way to essential architectural principles of future generations of mobile cellular networks. All such standards up to and including 5G are based on the major assumption that the network fabric provides communication services between end points (such as terminals, and cloud centres). Nevertheless, a fundamental change in the nature of applications is foreseen that requires a redesign of future networks.

This phenomenon arises from the growth of Cyber-Physical Systems (CPSs), which become widely present by seamlessly connecting individual systems into a unified cyber-physical continuum. Fundamentally, such applications consist of a feedback loop that draws state information about the real world towards a point of computation, where the state information is processed, eventually generating a new control output. A wide spectrum of such CPS applications is envisioned, which includes on the one hand Augmented Reality/ Virtual Reality (AR/VR), wearable robotics, and on the other hand applications from the industrial automation domain such as closed-loop control, robotic planning, and coordination, as well as safety systems. All these applications have the commonality that communication and computational services need to be provided by the networked infrastructure that is catered to the application in an individual yet scalable fashion.

In CPS applications of AR/VR and wearable robotics, computational offloading is also expected to provide solutions for achieving energy efficiency, wearing comfort, and reliability. However, their realization demands that the network infrastructure is highly deterministic at the end-to-end (E2E) level, as unpredictable connectivity would directly jeopardize the operation and user experience. Such an E2E deterministic network infrastructure needs to consider heterogeneous domains – i.e., wired, and wireless networking infrastructure, cloud, and application domains – and must achieve required levels of security.

The fundamental performance metric for all CPS applications is the E2E system responsiveness, from the input to the network infrastructure (providing the state information) down to the reception of the feedback from the networked infrastructure. Over the spectrum of CPS applications, wide ranges of requirements exist with respect to the responsiveness of the networked infrastructure, however, all CPS applications require responsiveness within certain minimum and maximum latency bounds with a high confidence. This constitutes the deterministic responsiveness of future networked infrastructures as the key characteristic that needs to be supported.

A deterministic service provisioning (which can be the individual communication or the compute, as well as an aggregated response, being a concatenation of communication and compute) is given by a system providing the service within a well-defined time period accompanied with a well-defined level of confidence of service provisioning. The narrower the time period (i.e., the smaller the difference between maximum and minimum time bound), as well as the higher the confidence level, the higher is the deterministic behaviour of the system to provide the service. The fundamental goal of DETERMINISTIC6G is to develop architectures that enable ubiquitous CPS provisioning in future networked infrastructures through enabling E2E determinism to the level of reliability required by the corresponding CPS application.

This necessitates several key ingredients to be defined and enabled in future systems:

(1) Requirements and computational footprints to networked infrastructures, as well as their requested level of reliability.

(2) Networked infrastructures will have to maintain a notion of confidence to support E2E deterministic communication at run-time towards the running applications. Due to the quantitative nature of this notion and various stochastic influences exposed by the networked infrastructures, the characteristic of predictability becomes central to the service provisioning. Predictability of service provisioning can be defined as the ability to bound the stochastic uncertainty of the service provisioning within a given time period and over a given prediction horizon.

(3) Given the required levels of reliability to provide services for a given set of applications, their communication and compute workloads, and a run-time notion of the confidence that the system is able to provide the required service, the final element of future network infrastructures is the ability to adapt the providable system reliability, the requested reliability, as well as the application workloads dynamically at run-time. For this, applications will have to provide different service and workload envelopes, together with lead time definitions for transiting among the envelopes that can be activated through a dynamic interface at run-time.

This system approach comes with completely novel challenges with respect to application definitions and design, service provisioning and run-time adaptation. In this document, we start with the identification of some of these challenges regarding a selected subset of CPS applications, namely towards the application areas of Extended Reality (XR), wearable robotics, and industrial automation.

1.2 Objective of the document

This document represents the first technical report towards the identification of the challenges to achieve an E2E deterministic network infrastructure for the domain of CPS applications. This report sets the ground for the development of new technical concepts that enhance the predictability and determinism of wireless communication. To this end, the report elaborates visionary use cases that can be further enabled and realized in their full potential with upcoming 6G technology. Key Performance Indicators (KPIs) and KVIs, which will be derived from these use cases, will help to develop various technical enablers of the next generation E2E deterministic network infrastructure, including 6G wireless communication standards and evolved deterministic communication standards (TSN, DetNet). These technical enablers together with identified key architectural principles will be the basis of DETERMINISTIC6G architecture.

The aim of this document is to provide fundamental inputs from the application perspective towards the technical work packages. A set of possible promising use cases is presented that support the realization of current and upcoming trends in industry. The reader is introduced and familiarized with envisioned applications in the industrial context which are enabled by improved determinism in wireless communication. The descriptions are focusing on communication aspects of the use cases and present the challenges to realize them. Therefore, it is not intended to provide a fully-fledged definition of the use cases that might be directly translated into application requirements.

Besides the definition of use cases, it is the objective of this document to introduce central concepts, such as the KVIs, which are a novel approach of representing the societal values of technological advancements. Additionally, an overview of the areas of research and development, for which the project intends to provide contributions, is given. Figure 1.1 further depicts the structure of the document.

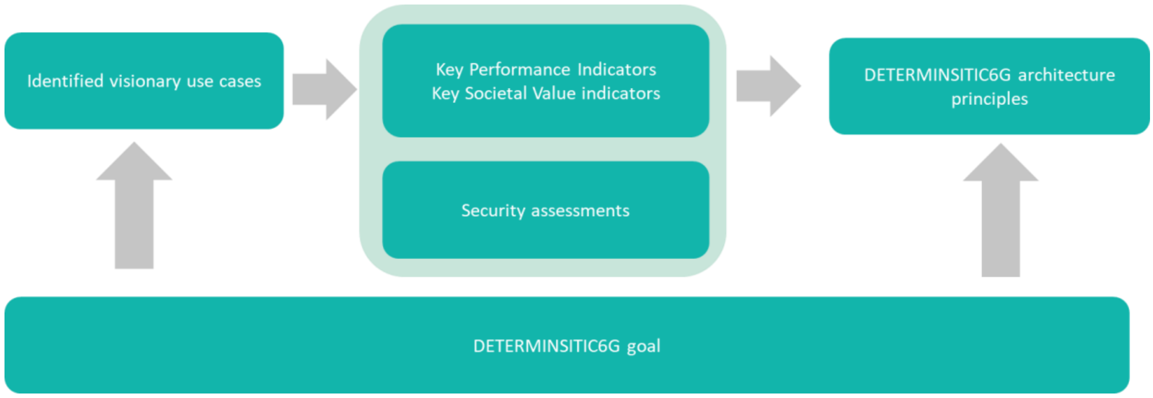


Figure 1.1: Structure of the document

1.3 Relation to other work packages and deliverables

Among the technical work packages (WPs), it is WP1 which intends to investigate the requirements for technical concepts and technological advancements. By looking into visionary 6G-based use cases, provisioning of novel architectural aspects, definition of social value indicators, and performing a focussed security analysis, the basis for further work in the DETERMINISTIC6G project is set. Especially, the findings of this first technical deliverable D1.1 act as an input for other WPs and subsequent deliverables in WP1. Figure 1.2 illustrates the relation between the WPs and their technical focus in DETERMINISTIC6G.

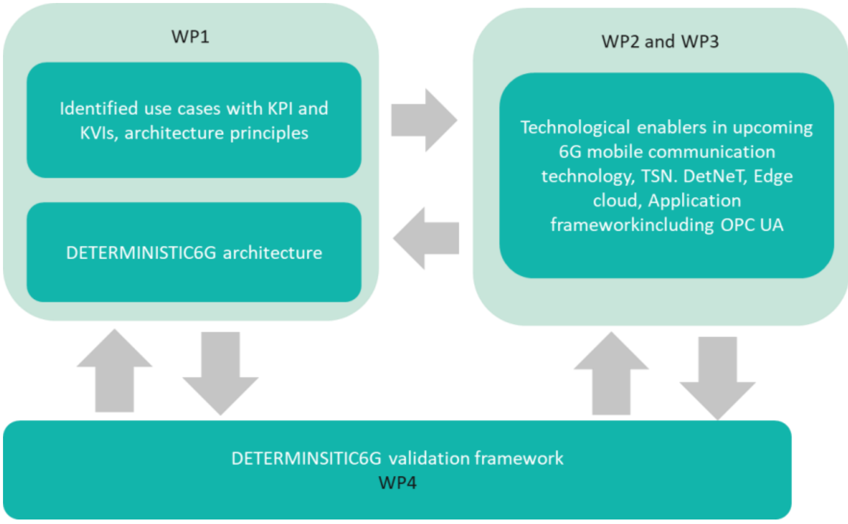


Figure 1.2: Relation between work packages in DETERMINISTIC6G

The use cases defined in D1.1 are taken in WP2 to derive the requirements for the 6G-centric enablers for deterministic communication services and to drive the work on new concepts. Similarly, the contributions of WP3 for the 6G and wire-based converged networks for deterministic End-to-End (E2E) communication, including edge computing services, are based on the definitions and requirements provided in this document. In WP4, a validation framework for 6G deterministic E2E communication is created. Initial simulations using this framework will be based on the structure and communication requirements originating from this document. The output of all these WPs is finally fed back to WP1, where a deterministic service definition and the 6G deterministic E2E communication architecture are developed.

1.4 Structure of the document

After this introduction, the deliverable continues in Section 22 with the definition of terms, which are in the focus of the project and of which the reader should be made aware. The major content of the document is presented in Section 3, where four visionary use cases requiring deterministic E2E communication are described. Section 44 introduces the concept of societal key values and corresponding key value indicators. An outline is given, how these concepts can be brought into the context of the proposed use cases and how to proceed with the work on them during the project. The principles of security assessment, which will be followed during the project, are stated in Section 55. In Section 66, an initial elaboration of computing resource requirements in the context of the use cases is provided. Afterwards, the DETERMINISTIC6G methodology and architectural concepts to achieve the targeted deterministic E2E communication in 6G are discussed in Section 77. Finally, Section 88 concludes the document.

2 Terms and definitions

This section introduces a couple of elementary terms which shall be known to the reader of the document. These terms are defined based on the interpretation of the authors of the document and might be specified differently in another context.

Term	Definition	Notes
Aperiodic	Type of recurrence. In the context of communication relations, aperiodic recurrence refers to communication that happens infrequently and unpredictably. It is typically triggered by an event that is not scheduled in advance. For instance, the change of a physical value, activation of an emergency stop, or the de-/registration of autonomous vehicles on a shared network.	Sometimes also described as "*sporadic*" or "*event-based*" communication.
Communication Cardinality	The cardinality describes how many communication partners might be involved in the communication relation on the sender and receiver sides. For instance, 1:1 means that one sender is communicating to one receiver, while 1:N means that one sender is communicating to N receivers.	
Communication Duration	The communication duration denotes the time interval how long a stable communication relation needs to be maintained in order to fulfil the purpose of the communication. For instance, the communication relation between two collaborative autonomous vehicles only needs to be maintained for the duration of the collaborative task.	
Dependability	In the context of real-time communication, dependability refers to the ability of a system or software to consistently deliver the expected functionality and performance while ensuring its correctness and reliability. It encompasses several key attributes that are crucial for real-time systems, including availability, reliability (see below), safety, fault tolerance, timeliness, and predictability.	Simplified, a system can be declared as dependable, when it is able to provide the required functionality and performance and can convincingly prove that it is able to do so also in the future for the anticipated and agreed operational conditions.
Determinism	All events in future can be determined completely based on past events and the laws of nature.	Also, referred as causal determinism in philosophy.

Deterministic communication	Given a message, the correct transmission is guaranteed to be performed in a specified period of time (not faster, not slower).	
Latency	Latency describes the required time to send a packet from a given sender to a given receiver over a given network.	Sometimes also described as "*packet delay*" or "*network delay*".
Packet Delay Variation (PDV)	PDV describes the amount of variation of the latencies perceived when a series of messages is transmitted from a given sender to a given receiver over a given network.	Sometimes also described as "*jitter*".
Periodic	Type of recurrence. In the context of communication relations, periodic recurrence refers to communication that happens periodically (with a given time interval, e.g., 1ms). The temporal distance between successive communication events is expected to remain within defined bounds. Examples are the current position and speed of a vehicle or (industrial) drive for coordinated or synchronized movement.	
Predictability	Predictability is the ability to accurately predict parameters of stochastically evolving KPIs of a given communication system over a given time interval at run-time.	Predictability requires the uncertainty of the system to be bounded.
Prioritization of communication relationships	The corresponding value indicates the relative importance of the communication relation compared to other communication relations. It depends on multiple factors, which might even include local legislation or preferences of the system operator. For instance, the transmission of an emergency stop signal is more important than the video stream of a surveillance camera.	
Quasi-periodic	Type of recurrence. The nature of quasi-periodic communication is similar to periodic communication, where communication events happen based on a defined period. But in contrast to periodic communication, the temporal distance between successive communication events is subject to higher variations (e.g., due to varying processing overheads).	

	Examples are video streams, where the higher variations in packet arrival times may be tolerated.	
Recurrence	Recurrence describes the regularity of the appearance of an event. In the context of DETERMINISTIC6G, this is mainly communication events.	
Reliability	Reliability describes the probability that a system will meet its expected performance metrics and perform its intended functions consistently and correctly over time.	Typically, reliability is expressed in terms of the frequency and impact of failures, e.g., as Mean Time Between Failures (MTBF). In case of a communication system, the reliability could, for instance, be expressed as the percentage of packets that are correctly delivered.

3 Use cases

3.1 Introduction

Over the last decades, continuous advances in information technology have enabled innumerable applications and use cases that change the everyday life of billions of people around the world. Beside many others, these application areas include entertainment, the financial sector, transport, health care, or industrial production. Meanwhile, one could even argue that modern societies depend on the benefits provided by this technological domain.

For most of these usages, the reliable exchange of information is a paramount prerequisite. Especially, use cases including interactions with the physical environment may lose their value if the reliability of communication is compromised. One application domain, where this is obviously the case, is industrial automation. Dependable communication is required to implement closed-loop control with the physical world. Thereby, physical variables need to be measured, processed, monitored, and controlled in real-time in order to achieve the desired effect. At the same time, industrial automation has high demands on functional safety, and even cyber security is a topic that cannot be disregarded anymore.

Multiple economical, ecological, and societal trends are driving the need for advanced and innovative automation. For instance, providing solutions for the world food production, sustainability of manufacturing, individualization of products, as well as the conceivable labour shortage also create new challenges for the next generation industry. Improvements in the dependability and flexibility of communication technology are expected to be key enablers for new use cases that will contribute to a better future.

Within this section, several visionary use cases are presented that will benefit from improved determinism of end-to-end (E2E) communication in converged wireless and wired networks. The use case descriptions will focus on novel scenarios for which this determinism is a key enabler. Due to the visionary character of the use cases, the reader shall be aware that the real implementations of these use cases in the future may be significantly different from the content presented in this document.

Figure 3.1 illustrates the use case domains that will be covered in this document. The use case presentations start with the extended reality technology, which can be exploited in many different domains and usage scenarios. Beside others, also for the context of industrial applications and the future working environment a high potential of using extended reality is predicted. Thereafter, the use case description of exoskeletons in the industrial domain is provided. These devices are applied to support workers by reducing the burden of physically hard and/or repetitive work.

Figure 3.1: DETERMINISTIC6G use cases

In the domain of industrial automation multiple disciplines may be distinguished. *Industrial process control*, for instance, refers to the application of control equipment and control theory to manage complex, continuous industrial production processes (e.g., oil refining, pulp and paper production, or chemical processing). However, due to the less adaptive and more static nature of these processes, there is not an intensive need for wireless communication. *Machine automation* intends to minimize the extent of human interactions required to operate discrete machines for mass production. The trend towards smaller physical dimensions of machinery and often very challenging requirements on reaction times and motion synchronization make the application of wireless communication technology infeasible for many contemporary machines. In contrast, for *factory automation* multiple specialized machines, robots and IT equipment are integrated to form production lines. Optimization of productivity from ordering to delivery of products is of high importance. For achieving this goal, highly adaptive production lines with flexible transport systems do have an advantage over conventional settings. Finally, in the last decades, *mobile automation* has attracted significant interest with many innovative fields of use. Partially and fully autonomous vehicles interacting with a hardly predictable physical environment require novel technologies and solutions to make efficient and safe operations possible.

While use cases in all these disciplines will benefit from improved deterministic E2E communication in converged networks, the focus of the remaining sections in this section will be on those scenarios which can hardly be implemented with the required dependability when using current technology. The adaptive manufacturing trend in factory automation, as well as the smart farming use case from the mobile automation domain depict a prosperous future of the next generation of wireless communication technology.

3.2 Extended Reality (XR)

3.2.1 Overview

Extended Reality (XR) is an umbrella term that covers immersive technologies that create an environment containing real or virtual components or a combination thereof, where the variable X serves as a placeholder for any form of new environment (e.g., Augmented, Assisted, Mixed, Virtual, or Diminished Reality) [AKP+21]. XR has a vast range of applications in almost all industry segments, including health care, shopping, and manufacturing. In this document the focus is on XR in industrial scenarios, where use cases include for instance XR for teaching, training, visualization, collaboration or maintenance and control. In the industrial domain it is expected that XR will be used in connection with many other use cases as well, e.g., as a means of communication with machines. With surveys [ECL+22] showing that hybrid work is expected to be the norm with 40% working remotely and employees seeking flexibility, XR is foreseen to be an enabling technology supporting this trend. In this report we consider the following two variants of XR: VR, where the user is totally immersed in a simulated digital environment and AR, where digital information is overlaid on images of reality viewed through an AR device.

3.2.2 State of the art

With major ecosystem players starting to invest in XR, it is expected that over the coming years there will be a major shift in human communication from mobile phone devices to wearable glasses technologies [AKP+21]. While VR Head Mounted Devices (HMDs) have already started to reach scale, AR HMDs have not reached mass market adoption. However, midterm AR glasses are expected to take the lead. Depending on the virtual world to be created and the AR overlays to be realized, different degrees of immersiveness are possible in XR going from e.g., very basic text overlays and simple virtual worlds to real-time image overlays and complex virtual worlds that are highly immersive. While today's devices can meet the requirements imposed by limited XR experiences, they are still in need for improvement, and they are definitely not able to support the future need for fully immersive experiences.

The XR devices of today are typically tethered to a smartphone or a computer. Most of them are still quite bulky, not very lightweight, encounter issues with heat dissipation and still have limited attractive appearance. For broader adoption and reaching a mass market, XR devices need to become smaller, more lightweight, and fashionable. Early AR prototypes, e.g., Nreal glasses[1], are already going in this direction. At the same time more immersive experiences call for an increased need of processing power on the XR device.
Reducing the size and weight of the devices decreases the space available for battery and local compute on the device. To allow long battery life while preserving the quality of experience for the end user, compute functionality will need to be moved away from the device. This is referred to as compute offload. The level of compute offload can vary depending on the XR application and the capability of the device. The communication back and forth between the device and the edge/cloud can be realized over a wireless link. Depending on the functionality that is offloaded, different requirements are put on the communication link. The trend is towards an increasing level of offloading time-critical functionality, creating real-time sense-compute-actuate loops over multiple hops including the wireless network. For the realization of these, a network platform with integrated time-aware deterministic communication and computing is needed.

[1] https://www.xreal.com/

3.2.3 Use case description

Professional workers in the industrial domain already today face the challenge of an ever-evolving factory environment with new tools and machines being introduced, and processes becoming more and more complex. This trend is expected to continue. In 2021, a survey [ECL+22] revealed that around 30% of decision makers and production employees rank "difficult and time consuming to learn" and "complicated to use" as the most important drawbacks for tools. Both agree that the skills gap will grow as more tools are introduced. XR is expected to be instrumental in closing this gap. The possibility to have information overlaid on the real world while simultaneously having your hands free has been shown to increase worker efficiency dramatically. Furthermore, XR provides more possibilities for human-centric design, meeting the industrial worker at his or her individual development level, overlaying/showing only the content that is relevant for the individual worker.

Some examples of typical scenarios where XR can support the industrial worker are listed below:

- XR for teaching/training, maintenance/repair, e.g.,
    - Near-term: a worker uses AR to get instructions and support from a remote trainer/expert
    - Future: a holographic representation of the trainer demonstrates steps to the worker and provides immediate assistance
- XR for factory floor visualization and manipulation, e.g.,
    - Near-term: information about the status and next planned steps of machines is displayed on the workers XR device, while walking on the factory floor
    - Future: Full simulation and manipulation possibilities for the AR industrial worker making also use of the digital twin of the factory
- XR for collaborative design, e.g.,
    - Near-term: Several AR users working on a shared static digital model
    - Future: Several AR users working on a shared dynamic digital model, e.g., a model of a technical solution (e.g., tool, process, etc.) developed together that involves movement over time.

Use case system architecture

In order to meet the demands on XR devices while at the same time meeting the criteria on being more fashionable, with smaller form factor and higher user comfort, offloading parts of the XR processing to the mobile network edge is a recognized solution approach. Offloading compute to the network edge can in general reduce the device's complexity (cost, size, power consumption, and heat dissipation) and enable more powerful processing to be performed. Different offloading levels are possible, ranging from low offload where only few functionalities are moved to the edge to high offload where basically all functionalities are run from the edge server and only sensor data is sent from the XR device over the uplink. Studies have shown that low offload reduces the device's energy consumption by threefold while mid offload reduces it by fourfold and high offload can reduce the device's energy consumption by more than sevenfold [AKP+21]. A high offload does, however, also imply more traffic on the communication network including stricter requirements on bitrates and latencies.

Table 3.1 shows typical uplink (UL) and downlink (DL) bitrates as well as latencies and reliabilities for VR and AR as experienced today. A key limit here being that the user experiences motion sickness if images are not updated with sufficiently low latency. A common assumption is that rendering motion-to-photon latency (MTP) greater than 20ms (10ms) for VR(AR) starts to cause nausea when a human wears a VR(AR) HMD. By using compensation techniques such as asynchronous time warp [Wav16]

the end-to-end round trip latency requirements from headset to edge server and back to the headset can be relaxed to orders of magnitude of 60ms (supreme quality) to 90ms (basic quality).
XR traffic is characterized by a mixture of pose and video from/to the same XR device. The video frame size is varying over time and packets arrive quasi-periodically with application jitter after IP segmentation. Traffic arrival time to the Radio Access Network (RAN) is periodic with non-negligible jitter due to application-processing-time uncertainty. Video frame sizes are not fixed over time. Typically, packets arrive in bursts that must be handled together to meet stringent bounded latency requirements. The combination of large packet-size variance, arrival jitter, together with latency and reliability requirements present a significant challenge for today's mobile networks.

Table 3.1: XR requirements on video frame level

Application	DL bitrates (Mbps)	UL bitrates (Mbps)	One-way latency (ms)	Frame reliability (%)
VR	30-100	< 2	5-20	≥ 99
AR	2-60	2-20	5-50	≥ 99

The XR connectivity requirements depend heavily on the architecture, the level of offloading anticipated and the targeted Quality of Experience (QoE). For the future, the range of bit rates and bounded latency requirements is expected to increase even further compared to Table 3.1. For instance, higher DL rates are expected to accommodate for the request for higher resolutions and more advanced graphics, while higher UL rates are expected to accommodate for the request for more dynamic interactions as well as multi-user interaction. The rates can vary significantly over time depending for instance on the actions and movement of the user, the complexity of the environment, the complexity of the overlay, etc.
Figure 3.2 shows an architecture of an XR device. Here, two main functional areas are highlighted that benefit from offloading. These are media rendering and spatial compute. Figure 3.3 shows the architecture when these are offloaded towards an edge/cloud.
Media rendering:

- Rendering is the process of generating media (e.g., immersive video) to be shown to the user in the HMD. It depends computationally on the complexity and resolution of the graphics engines. In remote rendering, the application is rendered at a remote server, and then sent back to the HMD via the wireless network. This allows for more complex rendering, which is needed to achieve a more immersive experience, a higher resolution, frame rate, and more advanced graphics.
- Offloading media rendering results in high throughput **downlink traffic** with low bounded round-trip-time, thereby increasing the requirements on the downlink traffic.
- Decoding of immersive media can be executed at the server, and regular video (2D per eye) can be sent to devices, where legacy video codecs can be used for decoding on the device.

Spatial compute:

- Spatial compute functionalities are a collection of functionalities that are primarily used to gain an understanding of the local environment. Offloading spatial compute allows for the creation of more dynamic environments, more dynamic interactions as well as multi-user interaction and input.
- Offloading spatial compute results in high throughput **uplink traffic** with low bounded round-trip-time, thereby increasing the requirements on the uplink traffic.

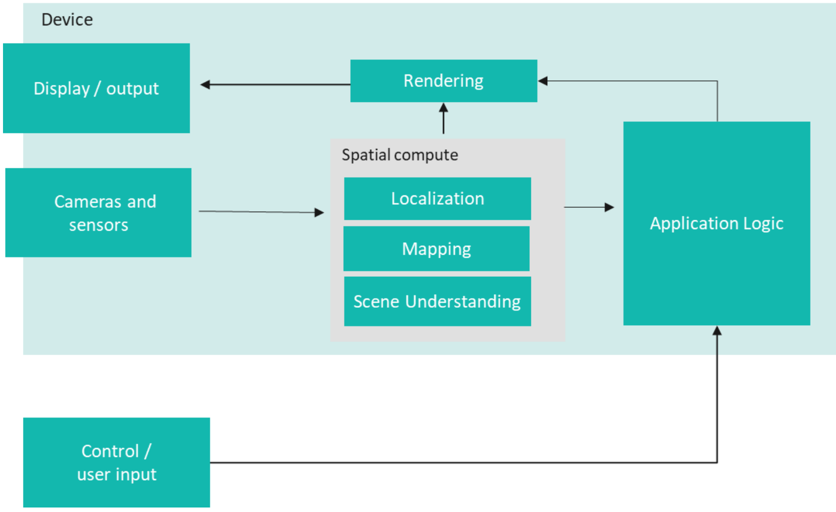


Figure 3.2: XR architecture, all functionality on the device

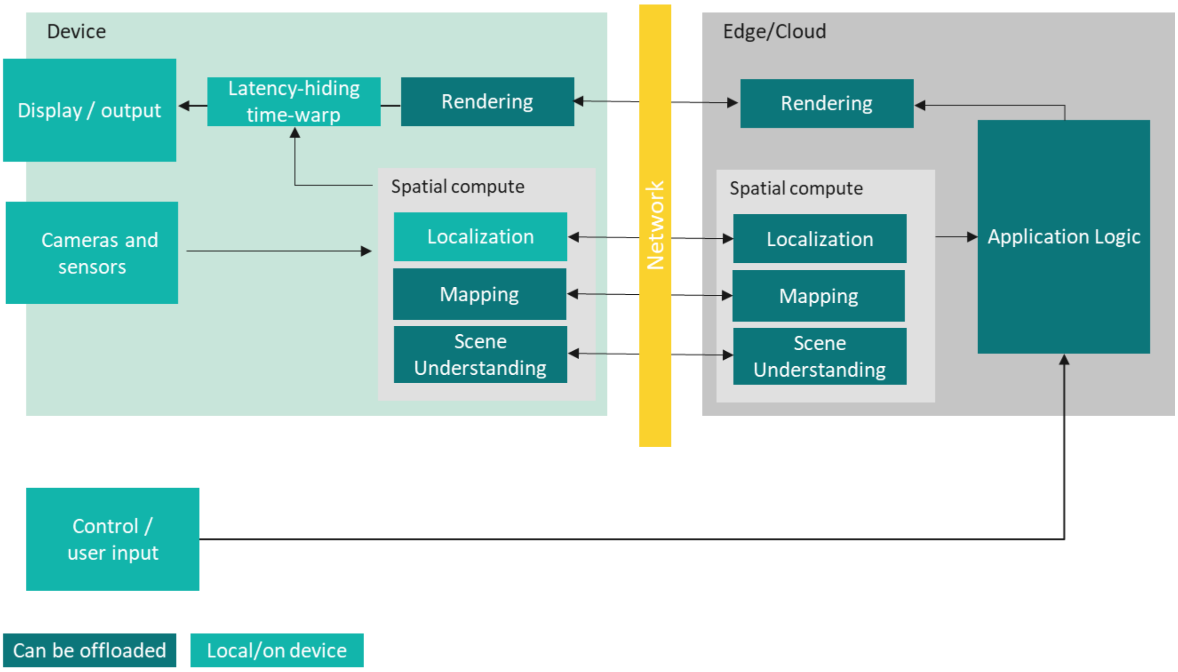


Figure 3.3: XR architecture, possibilities for offloading functionality

3.2.4 Communication relations and requirements

Fehler! Verweisquelle konnte nicht gefunden werden. shows the relevant communication relations and requirements for the use case of collaborative human-robot interaction using AR. The requirements depend on the scenario at hand and are only an example. Here, a use case is considered, where a worker is walking on the factory floor, retrieving information on the status of robots. The information is displayed as an AR overlay and used by the worker for manipulating the robots. When compensation techniques are used, packet losses can be tolerated to some extent but the degree to which this can be tolerated depends on the interaction and dynamicity of the human and environment.

Table 3.2: Communication relations and requirements for the use case of collaborative human-robot interaction using XR

Communication content	Communication purpose	Recurrence	Bandwidth (MBits/s)	Communication cardinality	Max Transmission Latency/ Packet Delay (ms)	Min Transmission Latency/ Packet Delay (ms)	Transmission Jitter/ Packet Delay Variation (ms)	Distance/ Position/ Velocity
Status reporting for robots/interacting devices	Reporting of change of value of monitored parameters (including motion position and speed of the robots)	Aperiodic	<1	N:1, from robots towards the smart factory cloud based industrial application	Not relevant	Not relevant	Not applicable	N/A
Video feed from XR headset	Update video image from user, Track user's actions	Aperiodic or quasi-periodic	UL: <2 (VR) UL: 2-20 (AR)	N:1, from mounted XR devices on the worker towards smart factory cloud based industrial applications	20	5	Depends on the application	Human walking speed
Video feed to XR headset	Update video image for user identification and position of the object (moving robots)	Aperiodic or quasi-periodic	DL: 2-60 (AR) DL: 30-100 (VR)	1:1, information dedicated to single human worker to be visualized on XR device	50	5	Depends on the application	Human walking speed
Control/user input from handheld device	E.g., confirm, interact, etc.	Aperiodic	<1	N:1 commands from N devices towards smart factory based industrial application	Not relevant	Not relevant	Not applicable	Human walking speed

3.2.5 Novel challenges

With a demand on even higher resolutions and more dynamicity for immersive experiences via XR in the future, the challenge to deliver all data within the latency bounds continues to be of highest relevance and requires new solutions. XR use cases in manufacturing do not have constant requirements over time neither on the network nor the computing resources but instead requirements vary over time and depend heavily on the dynamicity and immersiveness of the scenario at hand. A simple text overlay can be realized without high resolution, inspecting a quickly rotating part may require high dynamicity, etc. This makes XR applications highly relevant in the context of the cyber-physical continuum. Gaining a more fine-grained understanding of the communication and computational footprint of XR applications required for different scenarios and revealing this information to the network infrastructure, opens up for the possibility for the network to adaptively respond to the needs at run-time. Different service and workload envelopes could for instance correspond to different requirements on resolution, dynamicity, and computational workload. DETERMINISTIC6G will investigate how deterministic service provisioning in future network infrastructure can help in enabling end-to-end determinism and thereby support to deliver excellent Quality of Experience of XR applications in the future.

Furthermore, to realize XR at scale there is a need to provide the required service to many concurrent users in the network; for this RAN enhancements and network densifications are expected to be needed. Also, the XR services will need to co-exist with other types of services, as such service differentiation is needed and can be realized via network slicing and QoS related features. In a factory environment with a dedicated network, XR services may need to co-exist with other low-latency services such as cloud robotics and machine control. Progress on wireless deterministic networking will be highly relevant to the evolution of XR services. How to effectively meet the high uplink requirements for AR in combination with the stringent latency requirements at all times is a key challenge. Furthermore, XR will require an adaptive and optimized design with respect to privacy and security that takes into consideration both the data integrity of the XR users, and the persons and environment in the field-of-view of the XR users.

Figure 3.4 is taken from [ADF+23] and shows on a timescale the different envisioned operating points of XR together with their requirements. These range from simple applications (that can be satisfied with Mobile Broadband (MBB) traffic) to basic apps focused on conversational AR to evolved apps enabling dynamic environments to advanced apps that enable highly dynamic environments and interaction. It can be seen how the requirements on the downlink (DL), uplink (UL) increase. This also applies to the latency, which in Figure 3.4 is focused on the network and is given as the sum of the latency for RAN (Radio Access Network) processing and core network processing.

The different operating points could also be interpreted as different service envelopes required for different scenarios (conversational AR, dynamic AR with simple interactions, AR in highly dynamic environment and dynamic interactions) that a worker could encounter on the factory floor.

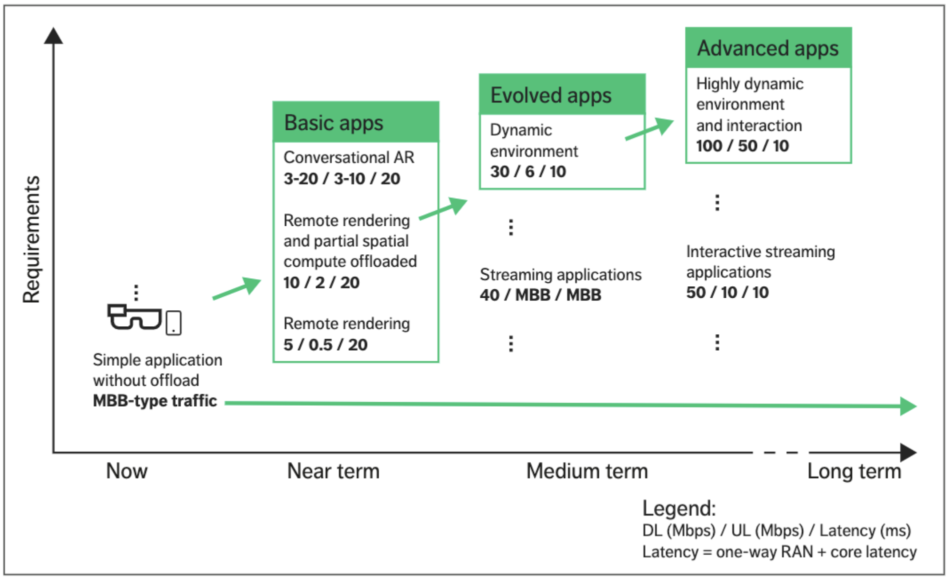


Figure 3.4: Evolution of requirements over time [ADF+23]

3.2.6 Security requirements

Besides security issues that are common to all IT systems, the use of mobile networks for XR applications involves specific vulnerabilities that need to be considered. Depending on the criticality of the application, i.e., the impact of deteriorated XR service on human or environmental assets, dysfunction due to unintentional or intentional actions needs to be rapidly detected and corrected.

Even though for some XR applications, such as smart manufacturing, a closed environment and private cloud computing is preferred, the systems are becoming more and more open involving non-managed devices, allowing accesses from the web, and using wireless communications (e.g., 5G and IoT networks) that introduce new vulnerabilities when compared to a proprietary closed system.

The vulnerabilities include for instance the following:

- 5G networks allow external applications such as XR to request edge computing or network resources to be allocated for own processing. This opens the possibility for deploying malicious applications in the edge computing platform within the 5G network and reducing the protection of sensitive information being offloaded.
- Time synchronization between an XR device and the edge cloud processing can provoke security and safety issues. In particular, augmented reality is very sensitive to time synchronization attacks.
- The signalling traffic used by the 5G network might be blocked or altered, or interferences may be introduced by attackers which will affect the reliability of the system.

The security requirements include for instance the following:

- Assuring the confidentiality, integrity, and availability of the XR use cases depending on the criticality of the use case.
- Dedicated security functions applied to safety in industrial processes (e.g., virtual emergency switch in XR, timely availability and prioritisation of crucial information in XR devices).
- Interoperability of security requirements in factories that involve different stakeholders.

- Considering the impact of Cyber Physical Systems (CPS) on communication, computation, and security and vice versa.
- Root cause analysis and forensics in case of accidents.
- Network, application and/or system behaviour analysis and anomaly detection

3.3 Exoskeleton in industrial context

3.3.1 Overview

Occupational exoskeletons are becoming effective tools that companies are adopting to reduce the physical burden of workers performing demanding activities. The envisioned use case is related to the deployment of exoskeletons in an industrial context which rely on future 6G networks for achieving a scenario in which exoskeletons can work in maximum synergy with workers, thanks to their capability of adapting their behaviour according to the user status, the type of task or other context-related aspects. The exoskeletons within the envisioned use case will become part of a network in which they exchange data with a centralized system and interoperate with different modules of the industrial ecosystem.

3.3.2 State of the art

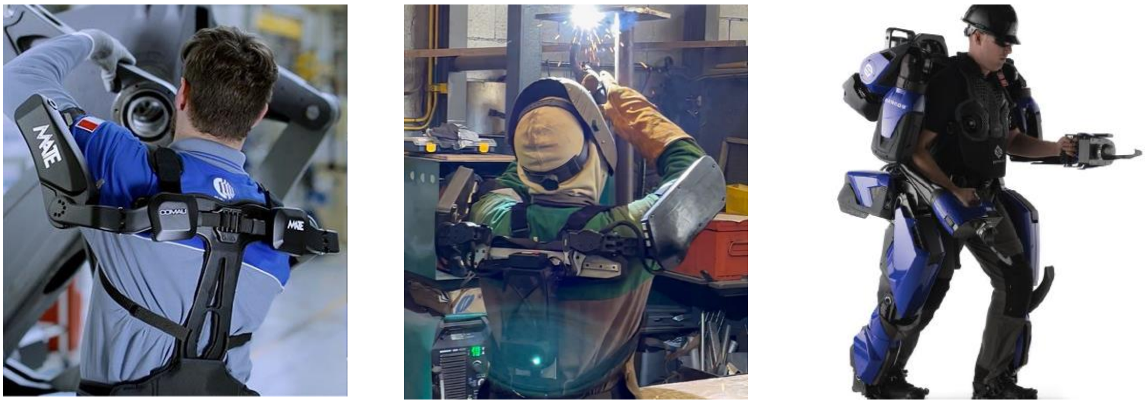
Based on the Sixth European Working Conditions Survey [EWCS15], nearly half of all workers in Europe experience work-related musculoskeletal disorders (WRMDs), which are the leading cause of occupational illness. The significant costs and health implications of these disorders place a substantial financial burden on both companies and healthcare systems [Par+17]. The majority of WRMDs stem from biomechanical overload resulting from repetitive overhead upper-limb movements [GD08] and the repetitive manual handling of loads and awkward postures. Recently, many companies have begun investing in occupational exoskeletons (OEs) to enhance the working conditions of their employees, with the expectation of reducing the incidence of WRMDs and improving productivity by lowering absence rates and enhancing worker well-being [HML+20] [MAD20].

OEs can be defined as personal wearable equipment that can help to reduce the physical load of workers performing demanding activities [MAD20]. OEs can be classified into two main categories based on the body part they target: lumbar OEs are aimed at reducing biomechanical overload at the spine level, while upper-limb OEs are designed to ease the strain of repetitive upper-limb movements [Cre21]. OEs architectures can also be distinguished based on their rigidity. If the OEs are composed of rigid kinematic structures that run parallel to the human body segments, they are considered rigid, while clothes-like suits that envelop body parts are referred to as soft [Cre21].

On the basis of their actuation principle, OEs can finally be classified in three subcategories: passive, semi-active or active exoskeletons [Cre21]. Passive OEs rely on spring-based mechanisms capable of storing and releasing energy in the different phases of the movement of the user. Some examples of upper-limb OEs are EksoVestTM (EksoBionics®, Richmond, CA, USA), the AirframeTM (Levitate Technologies®, San Diego, CA, USA), the ShoulderX (SuitX®, Emeryville, CA, USA), Paexo Shoulder (Ottobock®, Duderstadt, Germany), the Skel'Ex (Skel'Ex®, Rotterdam, Holland), and the MATE (COMAU, Torino, Italy). Examples of commercially available lumbar exoskeletons are the BackX (SuitX®, Emeryville, CA, USA), Laevo Exoskelet (Laevo, Rijwijk, The Netherlands), and Paexo Back (Ottobock®, Duderstadt, Germany). Semi-active exoskeletons are often an extension of passive systems, designed to adapt the passive behaviour of their spring-based mechanism by automatically varying the level of assistance or engaging/disengaging them through active clutches [GTP+20]. Active OEs are based on the presence of powered actuators for generating the assistive action of sensors and control units to monitor and synchronize robot action with the user's motion and of power supply units [dLBK+16].

Very recently, new types of exoskeletons have been proposed, such as hybrid active-passive exoskeletons, combining passive and active solutions at different joints [MLT+22].

Figure 3.5 compares the three categories of exoskeletons: passive, semi-active, and active. Each type offers varying levels of assistance in mobility and performance enhancement, with active exoskeletons providing the most advanced and customizable support.



a) b) c)

Figure 3.5: Examples of occupational exoskeletons. A) Passive – MATE (COMAU), b) Semi-active- H-Pulse (IUVO) c) Active – Guardian XO (Sarcos)

Semi-active devices, along with the active devices, have several advantages as:

- They can provide adaptive support on the basis of the user's or environment's inputs,
- they may lead to an increase of endurance, allowing users to perform physically demanding, tasks that would otherwise be impossible,
- they may be fully integrated with the smart factory digital ecosystem, allowing the possibility of a remote training, as well as a real-time monitoring of the system through the creation of a connected digital twin of the exoskeleton.

On the other hand, there are some disadvantages as:

- they can be expensive (depending on the number of powered actuators), making them inaccessible to many people who could benefit from them,
- they can be complex devices that require skilled operators to set up and operate,
- they may be heavy and cumbersome to wear for long periods of time,
- they require a power source to operate, which can limit their mobility and use in certain environments,
- they require real-time, deterministic networking of subsystems (e.g., on board sensors of OEs, wearable and smart sensors for monitoring the user's status and environment) and an elaboration through complex AI control strategies to deliver the correct amount of assistance depending on the user's needs.

Among several examples that can be found in the state-of-the-art two commercial devices (German Bionic - Cray X, COMAU - Mate-XT 4.0) represent two use cases similar to the one envisioned in the DETERMINISTIC6G project, being an active and a passive-sensorized exoskeleton that communicate with a central system to provide data about the devices usage and to perform analysis of the user's posture (information about wrong postures can be used to provide the user an alert or information for correcting it).

3.3.3 Use case description

During the project, a use case is investigated where a lumbar active OE, that is connected through the 6G network to an AI-based assistant, is able to collect information from different sensors and to determine task-oriented assistive strategies (Figure 3.6). Through the enabling technologies developed in the project, future AI tools can collect kinaesthetic data from the sensors embedded into the exoskeleton and identify, instant-by-instant, the optimal task-oriented assistive strategy. The AI tools will be aware of environmental information, e.g., the identifier of the lifted/handled load, and shall be able to optimize the intensity of the delivered supporting action, based on the specific task to be assisted. Furthermore, unfavourable loads, critical repetitions, risky movements, bad postures are immediately notified to the user with warnings.

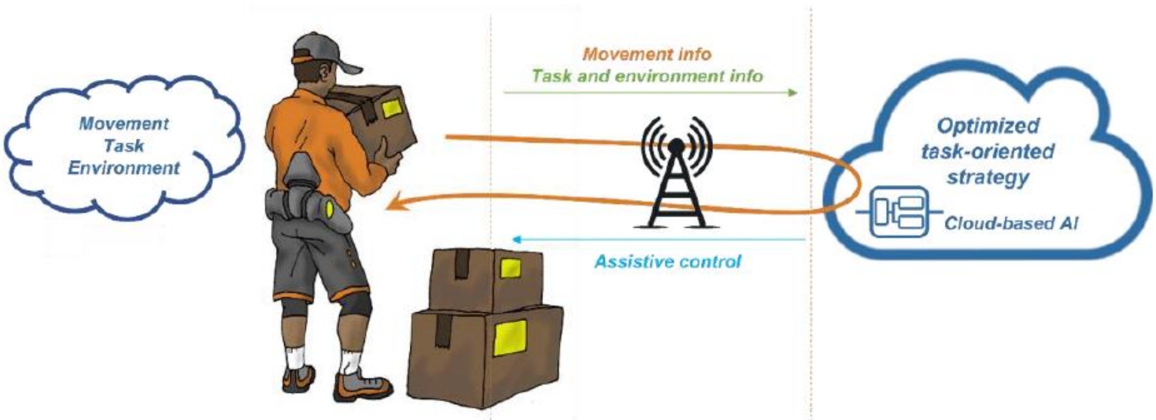


Figure 3.6: Lumbar active OE use case

The target user of the use case is a worker of a logistic warehouse. Figure 3.7 illustrates a typical user's action plan. He starts his day by wearing his lumbar exoskeleton, which is designed to support his back and reduce the risk of injury, allowing him to lift heavier loads without straining his muscles. Furthermore, since the exoskeleton provides mechanical power at hips level, it is also effective in providing walking assistance to reduce the effort of carrying loads while moving in the warehouse. The user then logs into the cloud-based system and the exoskeleton configurations customized for him are automatically downloaded. Besides this, the cloud-based system is connected to the warehouse's Enterprise Resource Planning (ERP) and shows him the list of shipments to organize and their priorities.

The user does his job aided by the support provided by the exoskeleton which adapts according to his movements and the load he is lifting. His work is also optimized by the updates on the status of each shipment which are provided by the cloud-based system.

In addition, a virtual replica of the system is created into the cloud-based system. This digital twin allows an external expert, such as an ergonomic specialist, to remotely monitor the worker's movements and provide real-time feedback and suggestions for posture correction. The expert can access the digital twin through the cloud-based system and analyse the worker's actions, including his lifting techniques, posture, and any potential areas of improvement.

By analysing the data gathered from the digital twin, the ergonomic specialist can identify potential risks and provide personalized recommendations to optimize the worker's posture and minimize the risk of injuries. Whenever the expert detects a suboptimal posture or movement pattern, they can send notifications or alerts to the worker through the cloud-based system, providing them with immediate feedback and guidance. The worker can then adjust their posture accordingly to ensure they are performing their tasks in the most ergonomic and safe manner.

Furthermore, the digital twin also serves as a valuable tool for training and development. The ergonomic specialist can use the data collected from multiple workers wearing the lumbar exoskeleton to identify common patterns and develop training programs to enhance overall performance and reduce the risk of workplace injuries.

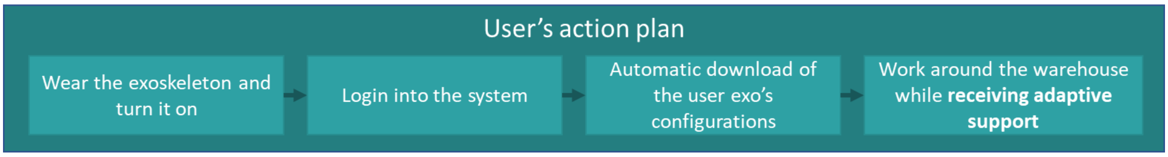


Figure 3.7: User's action plan for exoskeleton use case scenario

The use case aims to boost the maturity of active exoskeletons by leveraging the potential of the 6G network to offload the control of the robot, as well as the hardware components on-board the exoskeleton, thus reducing power consumption. A delocalized controller, monitoring, and real-time processing of a huge amount of information gathered from multiple subsystems will be made possible by enabling off-board complex AI assistive strategies, accounting for kinaesthetic, physiological, and environmental information.

Table 3.3 reports a brief summary of the 6G-enhanced exoskeleton main features and high-level requirements.

Table 3.3: Summary of 6G-enhanced active exoskeleton main features and high-level requirements

6G-enhanced active exoskeleton	**Main features**
	Adaptive support on the basis of the user's or environment's inputs.
	Full integration with the smart factory digital ecosystem.
	Low-demand power supply requirements on board.
	Low weight and bulkiness.
	High-level requirements
	Real-time, deterministic networking of subsystems (e.g., on board sensors of OEs, wearable and smart sensors for monitoring the user's status and environment).
	Elaboration through complex AI control strategies.

An active exoskeleton typically comprises several subsystems, whose specific design may vary. These subsystems usually consist of a frame or structure, which is the external support structure worn by the user and is often made of lightweight materials like aluminium, carbon fibre, or plastics. Electromechanical actuators, usually located at the exoskeleton's joints, generate torque and are controlled by an electronic system. Sensors detect the user's movements or the surrounding environment and they can include accelerometers, gyroscopes, and force sensors. A power source, such as a battery pack or a portable power source, powers the exoskeleton's electronic and electromechanical components.

The control system is an electronic embedded system that manages the exoskeleton's movement and behaviour, with software, microcontrollers, and other electronic components. The user interface provides a way for the user to interact with the exoskeleton, control its movements, and receive feedback through buttons, displays, or other input/output devices.

Specifically considering the control system, it can be schematized into three main levels as in Figure 3.8:

- *Low-level control:* refers to the basic control of its actuators, such as motors or hydraulics, that enable the exoskeleton's movement. It receives input from sensors embedded in the device and from the middle-level control (i.e., torque reference) and computes the precise instructions (i.e., current commands) to transmit to the actuators. This level of control is responsible for the fine movement of the exoskeleton, guaranteeing that it moves smoothly and accurately in sync with the user's actions.
- *Middle-level control:* refers to the part of the control responsible to translate the high-level commands into reference commands for the low-level control.
- *High-level control:* is responsible for the overall control of the exoskeleton, in particular for the task recognition and for the interaction with the environment. The high-level control system gives output to the middle-level control system and elaborates additional information from external sensors, such as cameras, to make decisions about the exoskeleton's mode of action.

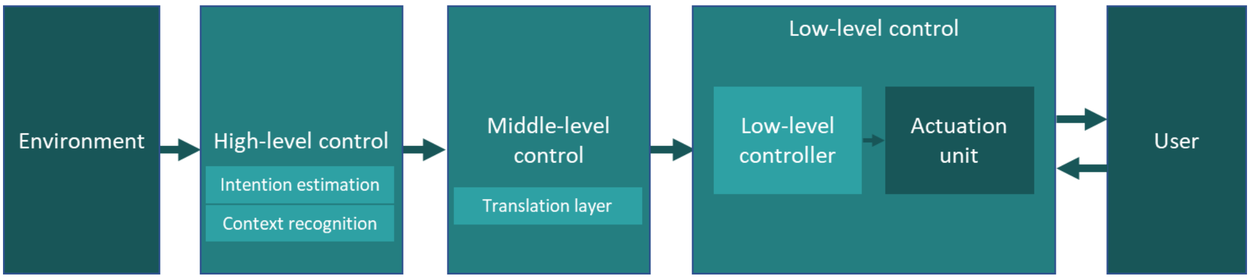


Figure 3.8: Hierarchical architecture of the exoskeleton control system.

Use case scenario

Depending on the computing offloading strategy, it is possible to imagine two different scenarios illustrated in Figure 3.9:

- **Near-term scenario**: Multiple exoskeletons sharing the perceived data acting independently. All sensors' data perceived by the devices are available on-line. An interactive dashboard allows remote monitoring of devices. The low-level control is embedded in the exoskeleton, while the middle and high-level control are delocalized in the cloud.

- **Long-term scenario:** Multiple exoskeletons sharing the perceived data with each other and solving tasks in a cooperative and/or coordinated manner. Targets of coordination can be several. It can be related to safety issues: as soon as the system detects the type of load, it can decide whether or not to involve (by alerting) a nearby colleague. It could also be related to the distribution of workloads within the warehouse. In this scenario all the control levels (low, medium and high) of the exoskeleton are delocalized in the cloud.

In both scenarios, a digital twin technology is employed to create a virtual replica of all the devices. The digital twins are constantly updated with real-time data from the sensors placed on the devices, equipment, and systems in the factory. The company has also partnered with a remote expert monitoring and intervention service provider. The provider has a team of experts who can remotely monitor the data from the digital twin and provide real-time insights to the factory operators and control the devices remotely (i.e., modifying the assistance provided by the exoskeleton as needed or providing real-time haptic feedback to the users through devices' actuators).

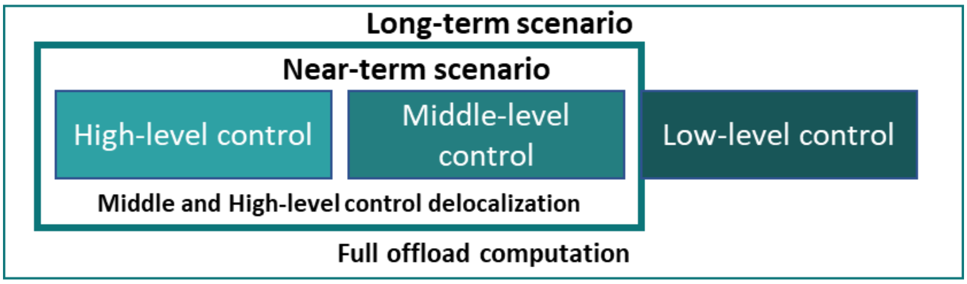


Figure 3.9: Offload computation perspective. Near term: partial control delocalization. Long-term: full offload of computation

Starting from a high level of abstraction, common to both scenarios, the architecture of the use case comprises three main subsystems: *the on-board system, the in-the-cloud system,* and *the smart factory,* as described below and as illustrated in Figure 3.10.

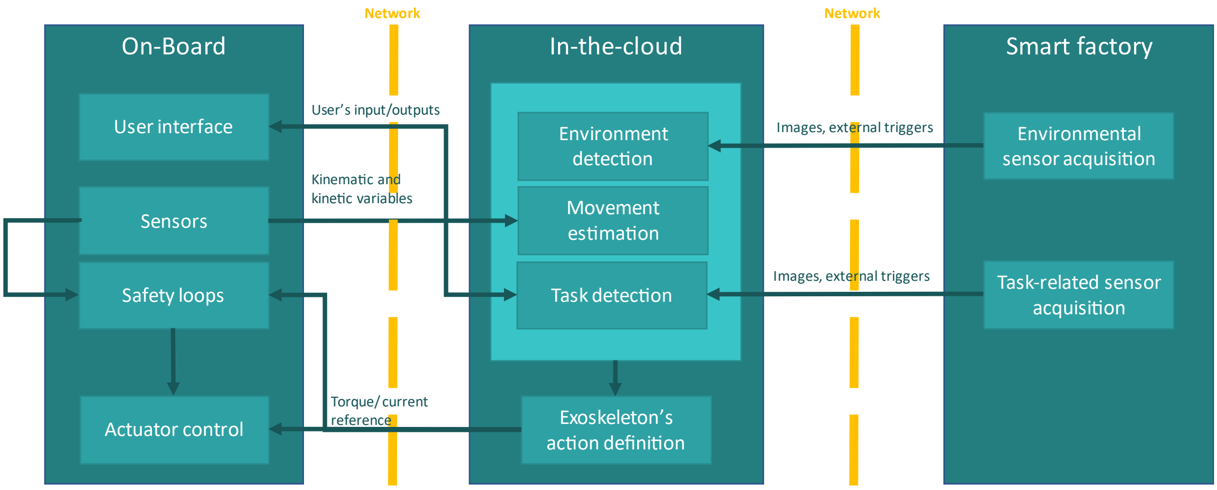


Figure 3.10: Exoskeleton use case main subsystems: on-board system, in-the-cloud system, and smart factory

On-board system

The on-board system includes all those components that must necessarily be physically incorporated into the exoskeleton: embedded sensors, actuators, motor drivers, and user interfaces (e.g., displays, buttons). Thus, the on-board system is in charge of embedded sensors data acquisition and motors control. In order to obtain a safe device for the user, it is necessary that all embedded sensors provide consistent readings and that the torque/current reference coming from the in-the-cloud system are compliant to the user's specific needs. For this reason, specific safety loops are designed to detect any faults and/or unexpected behaviours, preventing potential harm to the user.

In-the-cloud system

The in-the-cloud system is responsible of all the real-time operations required to define the exoskeleton's action, such as environment detection, movement estimation, and task detection.
It may also comprise services (e.g., company ERP) to access to company and task related information (e.g., daily shipments scheduling, inventory information).

Smart factory

The smart factory includes the external sensors (e.g., cameras) present in the relevant environment (i.e., the logistics warehouse) that can provide task-related information.

The on-board system and the in-the-cloud system communicate in a synchronous way: the on-board system provides to the in-the-cloud system information that is required as input by the control algorithm; on the other side, the in-the-cloud system communicates the output of the control algorithm to the on-board system.
Low latency and high reliability are two major requirements to allow that the torque provided by the exoskeleton is precisely synchronized with the movements of the user (i.e., the person wearing the device). In addition, they are two fundamental requirements in terms of safety since inputs required by the safety loops and their outputs rely on this communication.
Communication happens cyclically (i.e., at an established frequency), since it is crucial that the user keeps receiving support based on his/her movements: it is thus fundamental that (i) the in-the-cloud system regularly receives information related to the user's movements from the on-board system and (ii) the on-board system regularly receives information on the defined exoskeleton's action.

The in-the-cloud system and the smart factory communicate in an asynchronous way instead: the smart factory provides to the in-the-cloud system task-related information that is used as input for the definition of the exoskeleton's action.

Near-term scenario description

Figure 3.11 presents a focus of the near-term scenario main components.

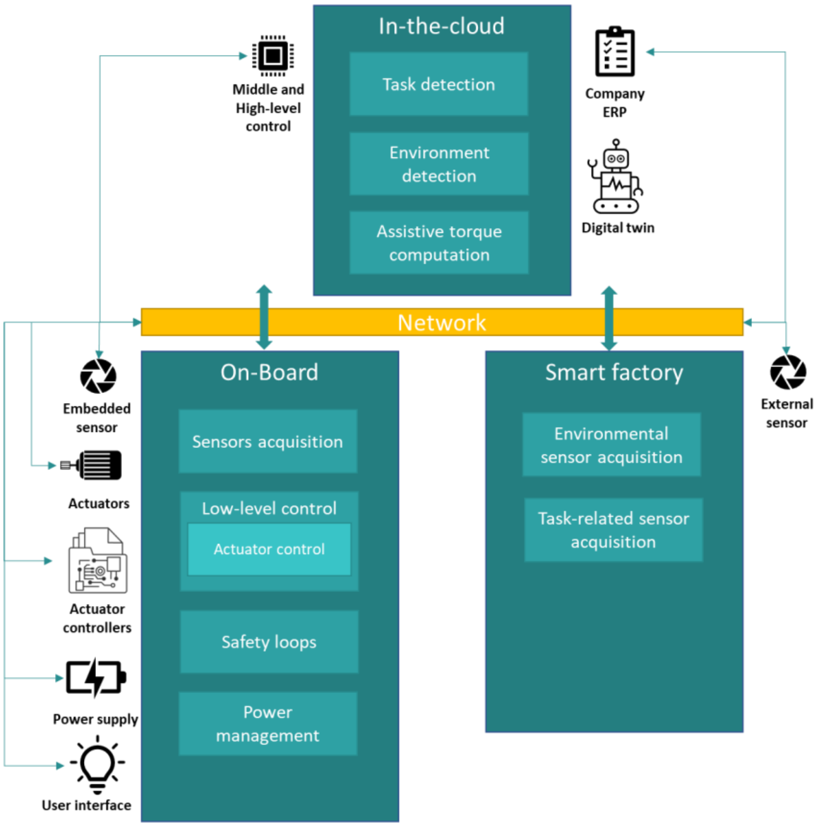


Figure 3.11: In-depth look at the near-term scenario main components.

In-the-cloud system

The in-the-cloud system is responsible of:

- Middle and high-level control algorithms (i.e., the assistive reference torque): These rely on data coming from the on-board system (i.e., embedded sensors, user interface, low-level control, etc.) and they are aimed to provide the commands to the low-level control on the on-board system
- Environmental detection algorithms: These take into account the inputs coming from external sensors (i.e., cameras, sensors on loads, etc.) to provide references to the high-level control and the task detection algorithm
- Task detection algorithms: These rely on information coming from the environmental detection algorithms and data coming from the on-board system to detect the user task and so, to give it as input to the high-level control algorithms that will compute the target support to the user, accordingly
- Digital twins: They are constantly updated with real-time data from both the on-board systems and the smart factory. They can be used by the remote experts to give insights to the high-level control algorithms.

On-board system

The on-board system is responsible of:

- Low-level control algorithms (i.e., the assistive reference current): These rely on data coming from the embedded sensors and from the in-cloud system (i.e., middle and high-level control outputs, etc.) and they are aimed to provide the current commands to the actuators

- Collect the signals from the embedded sensors
- Actively move the actuators. The current signal comes from the low-level control
- Power sources: The actuators, the embedded sensors, and all the electronics rely on batteries for power supply
- Safety algorithms: These rely on data coming from the embedded sensors and from the in-the-cloud system and they are aimed to control the status of the system, to avoid injuries to the user
- User interface: This is the main interaction point for the user with the entire network. It can request and send data to the in-the-cloud system (e.g., to the digital twin).

Smart factory

The smart factory is responsible of:

- The environmental sensor data acquisition.
- The task-related sensor data acquisition.
- Surveillance by the remote experts. They can interact with the digital twin in the in-the-cloud system to monitor the system and to give insights to the high-level control algorithms.

Long-term scenario description

Figure 3.12 presents a focus of the long-term scenario main components.

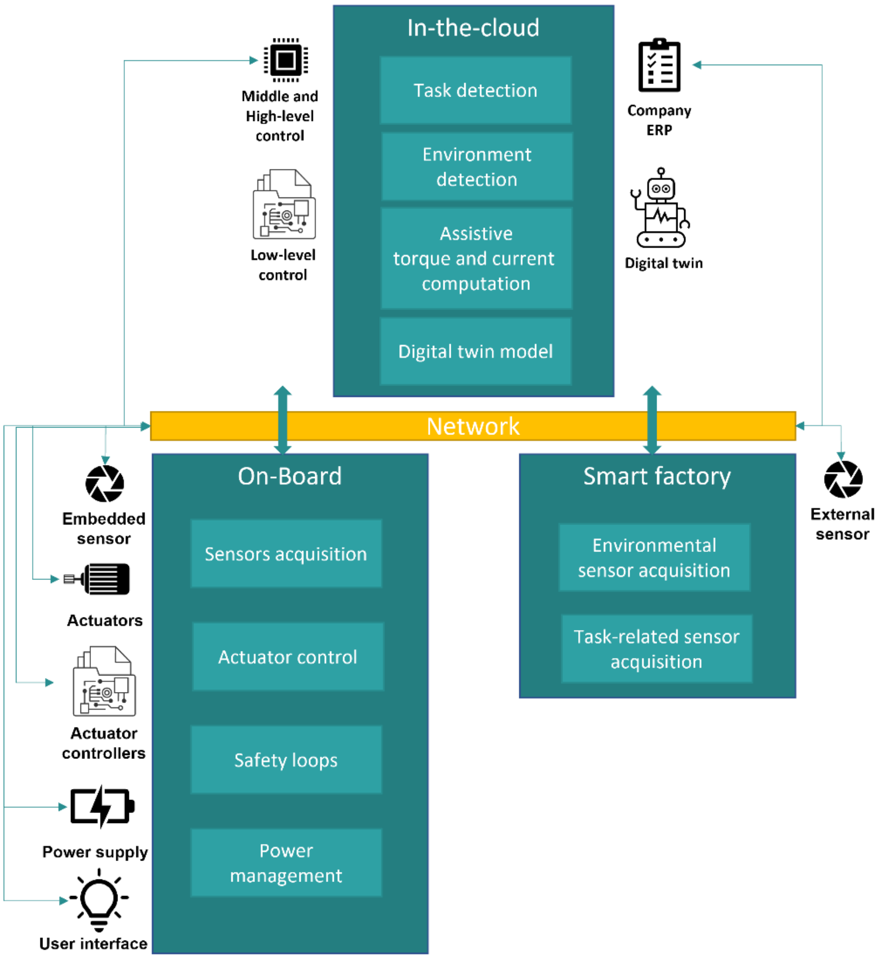


Figure 3.12: In-depth look at the long-term scenario main components.

As in the short-term scenario described above, it is possible to identify three main subsystems:

- in-the-cloud system,
- on-Board system,
- smart factory.

The long-term configuration differs from the near-term one since it foresees the offloading of the low-level control from the on-board system to the in-the-cloud system. Being the low-level control devoted to compute the values of the current to be provided to the actuators, its delocalization implies a higher frequency of data packets exchange between the in-the-cloud system and the on-board one, which in turn leads to more demanding network requirements (i.e., bandwidth, max latency allowed).

3.3.4 Communication relations and requirements

As described in the previous section, the architecture of the exoskeleton use case comprises different subsystems. The communication relations between them are described here: specifically, for each communication relation, the main communication content and its purpose are presented in a table, together with its corresponding KPIs.

The communication properties change between the Near-term scenario and the Long-term one: they are shown in Table 3.4 and Table 3.5, respectively.

The values entered in those tables refer to the presence of a single exoskeleton: in case of multiple exoskeletons, the bandwidth of the in-the-cloud system communication would increase, both uplink and downlink. In those tables the communication between each of the subsystems described above is detailed.

Table 3.4: Communication properties of the Near-term scenario. In the communication cardinality column, N represents the number of external sensors (i.e., cameras and switch/sensors on loads)

Communication content	Communication purpose	Recurrence	Period	Bandwidth (MBits/s)	Communication cardinality	Max Transmission Latency/ Packet Delay (ms)	Min Transmission Latency/ Packet Delay (ms)	Transmission Jitter/ Packet Delay Variation (ms)	Max Packet Loss per unit of time	Prioritization (of Communication Relationships) / Ranking	Communication Duration	Distance/ Position/ Velocity
Assistance to be provided	Deliver the proper assistance	Periodic	10 ms	0.1-1	1:1 , i.e., in-the-cloud system to on-board system	1.5	-	1.5	0	High priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
- Motors status - Info acquired by on board sensors - Battery level	- Monitor motor status/presence of alarms - Track user's movements - Monitor battery status	Periodic	10 ms	0.1-1	1:1, i.e., on-board system to in-the-cloud system	1.5	-	1.5	0	High priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
- Environment videos/images - Switch/sensors on loads	-Track user's actions inside the warehouse -Get info on lifted loads	Periodic, aperiodic	16 ms	0.1-30	N:1, i.e., smart factory to in-the-cloud system	50	-	50	1/16 ms	Low-mid priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Shipments status	Prioritize/ optimize user's work	Aperiodic	-	0.1-1	N:1, i.e., smart factory to on-board system	50	-	50	Not Relevant	Low priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s

The values entered in the table are based on the actual system architecture (e.g., data packets exchanged between middle/high-level control and the low-level control; typical frame rate of a security camera; distance and velocity considered respectively the maximum distance inside the warehouse - D_{max} – and the typical range of human walking velocity).

Table 3.5: Communication properties of the Long-term scenario. In the communication cardinality column, N_1 represents the number of cameras in the warehouse, while N_2 represents the number of switch/sensors on loads.

Communication content	Communication purpose	Recurrence	Period	Bandwidth (MBits/s)	Communication cardinality	Max Transmission Latency/ Packet Delay (ms)	Min Transmission Latency/ Packet Delay (ms)	Transmission Jitter/ Packet Delay Variation (ms)	Max Packet Loss per unit of time	Prioritization (of Communication Relationships) / Ranking	Communication Duration	Distance/ Position/ Velocity
Assistance to be provided	Deliver the proper assistance	Periodic	1 ms	1-10	1:1 , i.e., Middle/High-level control to actuators controllers	0.15	-	0.15	0	High priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Motors status	-Monitor motor status/presence of alarms	Periodic	1 ms	0.45-4.5	1:1, i.e., Actuators controllers to Middle/High-level control	0.15	-	0.15	0	High priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Info acquired by on board sensors	-Track user's movements	Periodic	1 ms	0.45-4.5	1:1, i.e., Onboard sensors to Middle/High-level control	0.15	-	0.15	0	High priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Battery level	-Monitor battery status	Periodic	10 ms	0.001-0.01	1:1, i.e., Power management system to Middle/High-level control	50	-	50	100	Low priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Environment videos/images	-Track user's actions inside the warehouse	Periodic	16 ms	0.1-30	N_1:1, i.e., External cameras to Middle/High-level control	50	-	50	1	Low-mid priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Switch/sensors on loads	-Get info on lifted loads	Aperiodic	-	0.1-1	N_2:1, i.e., External sensors to Middle/High-level control	50	-	50	Not relevant	Low-mid priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s
Shipments status	Prioritize/optimize user's work	Aperiodic	-	0.1-1	1:1, i.e., Company ERP to on-board user interface	>50	10	40	Not relevant	Low priority	Continuous (i.e., 100% of the time)	Distance = D_{max} Velocity = 0-2 m/s

In this scenario, the requirements table takes into consideration all individual elements that exchange data on the network, where the architecture can be likened to a flat network.

On top of these two scenarios, it is worth mentioning that in the event of a change in network performance, it is possible to imagine two modes of operation for the device.

- *Standard Mode*: This mode represents the default and standard operation of the exoskeleton. It utilizes the full set of sensors, both embedded and external, to gather information about the user and the surrounding environment. The control system optimizes the interaction torque and the target assistance level based on these diverse inputs, providing adaptability and responsiveness tailored to the user's needs.
- *Safe Operation Mode*: This mode is designed to prioritize user safety and maintain basic functionality even in situations where the network performance is compromised. In this mode, the exoskeleton relies primarily on the highest-priority data (i.e., the encoder readings from the embedded sensors used to guarantee a safe operation measuring the interaction torques with the user). Other non-essential information and external sensor inputs are temporarily suspended. The exoskeleton operates in a transparent manner (i.e., not providing assistance), zeroing the interaction torque with the user to prevent any inadvertent movements or actions. This mode ensures a safe and reliable operation until network performance is restored.

In order to guarantee an acceptable user experience, the switch between modes should be designed to be effortless and efficient without any noticeable gaps, delays, or inconsistencies in the functioning of the device.

3.3.5 Novel challenges

Occupational exoskeletons can reduce the physical load on the human worker. However, the full potential of OEs in the industrial environment is not being exploited yet, with only passive (i.e., not actuated) devices reaching the market and used in-field. Semi-active or active (i.e., powered) OEs offer wider adaptivity with respect to passive OEs, by automatically setting and adapting the level of support, according to the intended action of the users. Nevertheless, to ensure smooth interworking of powered OE with humans, demanding requirements are placed on computation and dedicated hardware, which result in higher weight and encumbrance than the passive counterparts and, consequently, poor acceptability by workers. Computational offloading offers significantly more efficient implementations of powered OEs in terms of energy efficiency and costs, accelerating its adoption across various segments of society, including manufacturing.

This offloading necessitates a new type of deterministic communication behaviour, supporting communications from the sensor through the cyber-physical representation to the actuator. Connectivity paths through the networked infrastructure with unpredictable latency variations directly jeopardize the operations of such applications.

To this end, a transparent collaboration between the application and the networked infrastructure would be crucial for achieving seamless integration and optimal performance of the system. This requires novel approaches to accurately capture and communicate the exoskeleton system's requirements, considering the unique characteristics and constraints of the offload computation. Likewise, the exoskeleton should receive updated information about the network status, to switch to the Safe Operation Mode when the communication requirements are not met.

In this regard, as analysed in the description of the use case, the application must be defined in terms of:

- data transfer needs,
- latency constraints,
- bandwidth considerations.

Additionally, being a time-critical application, it must provide acceptable communication latency, jitter, and data transmission reliability.

The networked infrastructure must also ensure end-to-end deterministic communication with the running exoskeleton application in real-time. Predictability, in this context, should mean that the networked infrastructure needs to reliably and consistently deliver communication services (i.e., target assistance, sensors data) to the exoskeleton, meeting predefined performance criteria (i.e., jitter constraints) and to real-time monitor any changes in the service performances. This would be crucial to accomplish safety requirements since all the computed actions must be tightly coordinated with the user's movements and intentions to avoid any unexpected behaviours or accidents.

Although multiple operation modes are not foreseen and described within the use case, apart from the one defined as the safe operation mode, it would be essential that the network infrastructure will be able to adapt its behaviour on the basis of network performances without any abrupt interruption of the service. This includes dynamically modifying communication protocols, computational resources allocation, and latency control to maintain optimal performance.

Based on these premises, deterministic wireless communication with 6G can significantly spark technological advancement of active OEs. A delocalized controller, monitoring, and real-time processing of a huge amount of information gathered from multiple subsystems, will boost the maturity of active OEs by enabling off-board feasibility of complex AI assistive strategies, accounting for kinaesthetic, physiological, and environmental information.

3.3.6 Security requirements

There are several security aspects that must be considered to ensure the safety and privacy of the user. Here are some of the key considerations:

- *Encryption*: All communications between the exoskeleton and the control system should be encrypted to prevent unauthorized access.
- *Data privacy*: The exoskeleton may collect and transmit sensitive data about the user, such as their biometric data, medical history, or personal information. This data must be encrypted and stored securely to prevent unauthorized access and disclosure.
- *Redundancy*: In case of a network failure or other disruption, the exoskeleton should have redundant control mechanisms in place to ensure that the user can still operate it safely.

Loss of time synchronisation, data loss and Denial of Service (DoS) should be considered as the most critical hazards since it can impact the correct assistive action provided to the user by the computation. The system has stringent latency requirements. Since it is in direct contact with the user, the communication of sensor data as well as the output of the control algorithms must be guaranteed to ensure the correct and not dangerous effect on the user.

The characteristics of the cloud environment for the 6G-enhanced exoskeleton are not yet defined. For safety concerns, as in the current implementation without computation offload, it could be reasonable to leave the system accessible through a local network.

In the short-term scenario, assuming the inclusion of interpolation and recognition logic within the in-the-cloud system, which allows for the distinction of artifact packets and/or the reconstruction of missing sensor data, the communication from in-the-cloud to the on-board system (high-middle-level to low-level control) can be identified as the most critical part since the data relates to the control of the exoskeleton (e.g., sensor data from on-board system to in-the-cloud, control output from in-the-cloud to on-board system). Less critical communications are those between smart factory and in-the-cloud system.

In the long-term scenario, a minimization of edge computing is foreseen in favour of greater computational load on the cloud. This will make it necessary to adapt the security mechanisms to deal with the changes in the vulnerabilities and risks inherent to each technology.

The impact of anomalies or security breaches is the instability of the system which will impact the assistance given to the user. Malicious insider or external actions can profit from a system that is not sufficiently protected by giving access to change user parameters or change the desired assistance. There are stringent latency requirements. The stability of the system, as it is in direct contact with the user, must be guaranteed.

3.4 Factory Automation: Adaptive Manufacturing

3.4.1 Overview

Adaptive manufacturing, an emerging trend in the field of industrial automation and production, has gained significant prominence over recent years. This innovative approach integrates advanced technologies, such as robotics, artificial intelligence, track-based or planar product transport systems, and data analytics, to dynamically adapt manufacturing processes on-demand to respond to fluctuations in market demand, product customization requirements, and supply chain disruptions, providing short changeover times, increased throughput on small footprint and maximum flexibility for new product configurations.

The following technologies are key enablers for adaptive manufacturing:

- Intelligent and flexible product transport systems: Individually controlled track-based or planar shuttles instead of e.g., conveyor belts, eliminate gaps in the output (e.g., caused by rejected products), increase throughput and are essential for on-the-fly changeover of reconfigured production lines, for instance, for new product configurations in small batch sizes.
- Robots and machine vision: In combination with machine vision, which detects properties like shape, size, orientation, and which is also capable of performing automated visual quality assessments of a product on a shuttle, robots provide highest performance and flexibility in pick-and-place operations.
- Digital twins and simulation: Digital twins and simulation eliminate the need for physical prototyping and enable feasibility analysis as well as optimization of the production process up-front with shortest changeover times.

The rapid escalation of global competition, coupled with the challenges posed by the COVID-19 pandemic, has accelerated the need for manufacturing agility and resilience, making adaptive manufacturing an indispensable strategy for organizations seeking to maintain a competitive edge in today's complex business landscape.

It is expected that the employment and integration of reliable 5G/6G technology will add a fourth pillar to the current key enablers, constituting a major contribution to production- and cost-efficiency of adaptive manufacturing, by e.g., the use of 5G/6G-controlled modular machine parts, collaborative Automated Guided Vehicles (AGVs) or optimized AGV routing.

3.4.2 State of the art

In the past, industrial manufacturing was based on production lines that consisted of heavy mechanical machines, and humans were an integral part of the production process. Technological improvements and increased productivity were driven by advancements in mechanical and electrical engineering. To achieve high throughput and thus lower the cost per unit, these production facilities were tailored to a certain type of product, which made them quite inflexible to changes. A product change typically requires a modification of the machine, resulting in undesirable downtime and cost-intensive interactions of skilled personnel. To keep productivity high, products were manufactured in large batches, and smaller batch sizes were often unaffordable.

The advent of computer systems introduced new possibilities in the automation of machines and production lines, enabling a high degree of flexibility. Modern industrial manufacturing involves complex interactions between highly automated production lines, IT systems (e.g., for order processing, production planning and monitoring, product shipping, etc.) and humans. A production line itself typically consists of multiple individual machines and components, where each of them is responsible for dedicated steps in the production process (e.g., unscrambling, filling, capping, labelling, quality control, packaging, etc.).

Initially, the control of machines was primarily based on the open-loop controller principle, which does not require precise sensing of physical variables. This reduces the complexity of control, but additional mechanical and electrical equipment is needed to maintain a stable flow of products. Therefore, even here the flexibility of production is limited since physical interconnections between production modules already define a product’s path through the production facility. Furthermore, a change of the product or its packaging often requires the manual adaptation of tooling equipment or even the exchange of whole production modules.

With the usage of closed-loop controllers the production process can be adapted dynamically in reaction to changing conditions in the physical world. This requires multiple sensors to observe physical parameters, a reliable transmission of the data to the controller, and the output of calculated setpoints by controllable actuators. Because most machines need to fulfil functional safety requirements and the productivity of the system shall not be compromised, the reliability of the control loop is of highest interest.

In the last couple of years, adaptive manufacturing has advanced to incorporate a range of cutting-edge technologies, such as machine learning, artificial intelligence, robotics, and advanced sensor technology. These technologies allow manufacturing processes to adjust in real-time to changes in demand, production issues, material availability, or product customization. The adoption of Industry 4.0 standards, characterized by the interconnection of machines and systems, has further enhanced the capacity for adaptive manufacturing in different industry segments, e.g., the medical device assembly segment [Pat23].

The term flexible automation or flexible manufacturing is often used in a similar context as adaptive manufacturing. Though both, adaptive and flexible manufacturing, aim at increasing production efficiency and versatility, they follow different principles and strategies. On the one hand, adaptive manufacturing employs advanced technologies like the Internet of Things (IoT), AI, and advanced sensors to monitor real-time production conditions and make instantaneous adjustments. It relies on real-time data collection and machine learning algorithms to interpret data, optimize operations, and reduce waste, making it highly dynamic and responsive to real-time changes in the production environment. On the other hand, flexible automation is built on programmable machines and robotics that can seamlessly switch between different tasks without significant downtime, thus supporting a high variety of different products or production volume on the same production line with minimal transition time. While flexible automation's versatility lies in the variety of tasks it can handle, adaptive manufacturing excels in its almost instantaneous adaptability to current manufacturing conditions [Mar23]. In the context of factory automation and the use cases described in this section, there will be no exclusive focus on the technologies that are essential for adaptive manufacturing, but rather a holistic point of view will be attempted, incorporating the key technologies of both adaptive and flexible automation.

The inherent flexibility of adaptive manufacturing calls for the application of wireless technology to further boost efficiency, reduce down-/changeover time and minimize human interaction. Exchanging tools on robots (and e.g., automatically delivering tools requiring maintenance to corresponding facilities or storage), automatically swapping modular parts of a production line to set up for a different product batch or optimizing production efficiency by enabling AGVs to move faster between production cells (or even further, by having them execute parts of the production process in movement) are use cases that require reliable and bounded low-latency communication between network participants.

Wireless communication in the industrial domain today to a large extent employ Wi-Fi technology. Most 5G installations in industry are pilots where the possibilities and limitations of the technology are explored. This was, for example, the case in the 5G-SMART project where 5G Release 15 was used for manufacturing use cases including mobile robotics[2]. Wireless communication presents several challenges when applied to adaptive manufacturing. Non-determinism in wireless communication can reduce the efficiency of a manufacturing process. Due to randomly evolving system variables like signal strength and interference, physical barriers, or a variable distance between sender and receiver, data transmission can experience varying latency, also known as packet delay variation (PDV). This can disrupt the timing of automated processes and lead to production errors or inefficiencies. One scenario, where this is especially true, are functional safety applications.

Functional safety applications trigger safe reactions and transition systems into a fail-safe state upon fault detection. The fault detection time includes signal propagation from a safe input to a safety controller, and the fault reaction time in turn includes signal propagation from the safety controller to safe devices in the machine to successfully execute the safety mechanism, e.g., stop the machine. Since the fail-safe state must also be triggered upon communication failure, corresponding timeouts on the communication layer must be set in such a manner, that a fail-safe state can still be reached when the timeout expires. In environments where data transmission is subject to varying latency, choosing timeouts at the lower boundary may lead to a high number of executions of the safety

[2] www.5gsmart.eu

mechanism (and thus system downtime), whereas choosing timeouts at the upper boundary will enable mostly uninterrupted system operation, but at the cost of lower performance (i.e., any devices or machine parts relevant for the safety application, may move only with a velocity that still allows establishment of the fail-safe state after expiration of the timeout). Reducing PDV allows for setting a lower (and more reliable) timeout, resulting in higher performance (e.g., increased velocity of movement) and less downtime.

The following sections expand on selected scenarios of the aforementioned use cases in detail, describe their communication characteristics and how they may be enabled by deterministic communication over 5G/6G networks.

3.4.3 Use case description

This use case for Factory Automation, and the more specific field of Adaptive Manufacturing, focuses on the following topics and main challenges that applications in these fields usually face and outlines different aspects which can be improved by the technological enhancements of the DETERMINISTIC6G project:

- Reliable communication between mobile and stationary devices
- Network participants competing for shared resources on the network, e.g., bandwidth
- Different communication requirements depending on the mode of operation

The reliable communication between different devices is the core requirement of any kind of Factory Automation application. While moving towards more flexible and more dynamic applications in the field of Adaptive Manufacturing, this type of reliable communication often cannot be provided by using wired connectivity anymore. This required shift towards wireless communication becomes prominent as soon as mobile devices are being integrated into Factory Automation applications. Such mobile devices / modules are often called Automated Guided Vehicles (AGV).

Using any type of multipurpose communication resource with multiple different devices always introduces a potential implicit competition regarding the limitations of the used communication and the available guarantees. These conflicts are also something that already had to be solved when looking at wired communication networks like OPC UA FX, or Time-Sensitive Networking (TSN) as, for example, the bandwidth of the network link or the employed functionality on the used infrastructure is limited. Shifting towards more flexibility via wireless communications aggravates this conflict as the number of communication partners accessing a single shared resource is expected to increase, in addition to the inherently random behaviour of the wireless communication medium itself.

Considering the general requirement for increased flexibility and the inherent conflict regarding shared resources (e.g., bandwidth) in a network with many participants, different applications in the Adaptive Manufacturing field introduce yet another challenge: The flexible nature of these applications regarding the extent to which they require certain network resources. An example may illustrate this best: An AGV following a pre-determined path on the shop floor will rely on some kind of obstacle detection to prevent collisions with human personnel or other AGVs. Using high-resolution image recognition is one possible technical solution, and while high-resolution images will likely provide the most accurate detection and object classification, lower resolution images may provide an equally reliable result, but with lower spatial precision or the calculations may require multiple subsequent images (which in turn may require the AGV to move at a lower velocity to be able to come to a complete stop before a collision). A similar trade-off would be necessary when comparing two modes of video streams with higher and lower frame rate. Additionally, there may be scenarios in which obstacle detection is not required at all, because other technical means prevent the possibility of any obstacle on the path, e.g., shop floor areas that are inaccessible to other AGVs or human personnel.

This example illustrates the highly dynamic nature of applications in the field of Adaptive Manufacturing, and the following subsections will elaborate more on this specific property, further called “operation mode” or “mode of operation”, referring to an application’s requirement to dynamically request network resources for limited duration with varying guarantees (e.g., bandwidth, max. bounded latency, etc., further referred to as “level of operation”), allowing the dynamic reallocation and more efficient usage of shared resources.

The following use case scenarios show different applications that showcase the main challenges described above.

UCS1: Cooperating AGVs

This use case scenario focuses on the cooperation of multiple AGVs with each other and with other stationary components of a factory. Such collaborative AGVs require tightly synchronized movement to transport heavy machine equipment, product parts, or participate in the production process while in movement.

The involved actors in this scenario are the following:

- AGVs
- Processing Cell
- Human personnel
- Safety System

Figure 3.13 shows these actors, possible movements and interactions as parts of the following application.

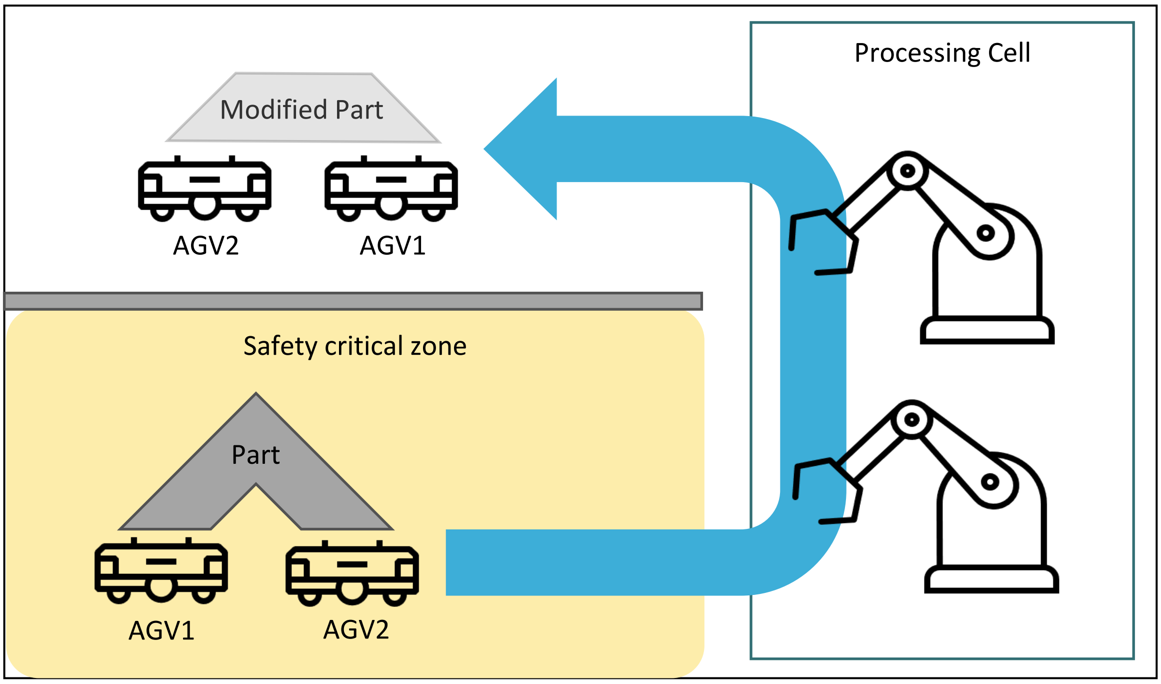


Figure 3.13: Cooperating AGVs

AGV1 and AGV2 are virtually coupled and move in unison towards the processing cell, transporting a given part. During this movement, the AGVs (further called "AGV swarm") traverse a safety-critical zone where an unexpected human interaction may occur. The precondition for this movement through the safety-critical zone is the previous registration to the underlying Safety System as soon as the AGVs entered the zone. When the AGV swarm leaves the safety-critical zone, it unregisters from the underlying Safety System.

After leaving this safety-critical zone, the AGV swarm enters an enclosed processing cell and registers itself to that processing cell. This cell is comprised of various stationary components performing different operations on the transported part. These operations can partly be done by the AGVs themselves and their coordinated movement relative to each other and the components of the processing cell. These movements could implement processing steps like bending, turning, inserting, or connecting individual or multiple parts.

The AGV swarm leaves the processing cell after all processing steps were executed successfully, transporting the modified part, or a newly produced part. The factory section following the processing cell is not accessible for human personnel, therefore, no safety considerations are required. This could be for example ensured by physical provisions, which are not relevant for this scenario.

These different operation modes also correlate with different communication relations, which are shown and described in the following paragraphs. The different communication properties are described in a generalized form, more detailed information and examples with value ranges can be found in Table 3.6. From the communication point of view, an additional role / actor next to the previously shown list was added: AGV swarm. The AGV swarm represents the combination of both AGV1 and AGV2. Other communication partners do not need to consider which specific device is responsible for the coordinated movement as this is abstracted by the AGV swarm. The swarm coordination role could be implemented by any one of the two AGVs or a separate process executing on additional resources like an edge computing or cloud instance.

Movement through the safety-critical zone:

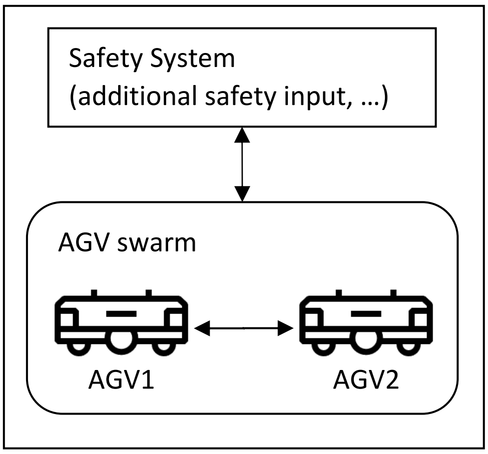


Figure 3.14: Communication between AGV swarm and Safety System

During this operation mode, while the AGV is moving through the safety critical zone, the following communication relations are of main interest.

Both AGV1 and AGV2 communicate with each other to implement the coordinated movement as an AGV swarm. The properties of this communication strongly depend on the implemented coordination logic. Usually, periodic exchange of control information is required. This communication is critical for the correct execution of the whole application and remains unchanged throughout all operation modes.

In addition to the periodic control traffic for the coordinated movement, communication is required between both the AGVs for any functional safety-related messages or events. The specific communication properties are again strongly dependent on the implementation, but most safety implementations use periodic communication.

Like the safety relevant communication between the AGVs, another safety related communication is required between the AGV swarm and the Safety System itself. This is required to communicate any safety relevant information from the AGV swarm to the Safety System itself or vice versa. This information includes safety-relevant events generated by the AGV swarm or any other connected system like light barriers or similar devices.

The final type of communication is the registration and deregistration of the AGV swarm to the safety system. This communication is purely event-based and does not require strict timing guarantees. The only requirement is that the communicated events do get delivered soon enough to match the current location of the AGV swarm.

Interaction with processing cell:

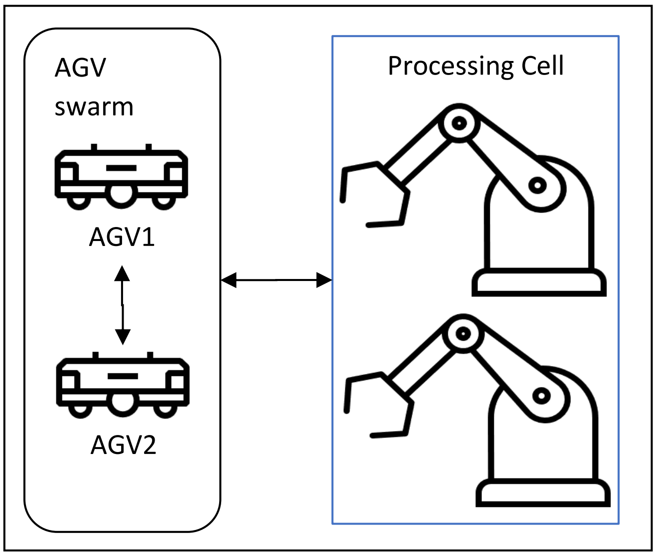


Figure 3.15: Communication between AGV swarm and processing cell

Like the previously described registration / deregistration communication with the Safety System, the AGV swarms needs to communicate with the Processing Cell or the underlying management system to register and deregister itself for the interaction required for the combined processing operation. This registration and deregistration communication is again aperiodic and does not require a strict timing or high bandwidth.

The communication within the AGV swarm stays the same in comparison to the previous mode of operation.

The most important communication relation for this operation mode is used for the combined processing operation and the associated movement of the AGV swarm relative to the processing cell. This communication has a high priority because it is essential for the correct behaviour and the intended processing operation. The coordinated movement is highly time-critical and any type of desynchronization between the AGV swarm and the processing call shall be avoided.

Transportation after processing:

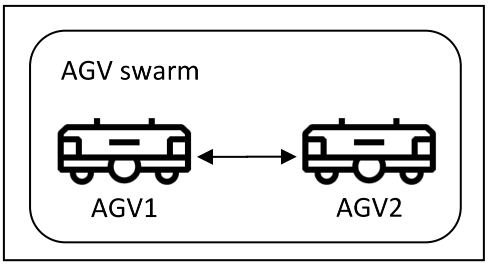


Figure 3.16: AGV swarm communication

The last mode of operation only requires the communication for the coordinated movement of the AGV swarm. This communication is unchanged in comparison to the previous modes of operation.

UCS2: AGV obstacle detection

This use case scenario focuses on the interaction of different independent devices within a shared physical region with limited shared communication resources. This scenario also shows the possibility of changing the operation mode depending on available resources. Any applications that support such flexibility thereby implement different levels of operation. These levels of operation and their potential usage are also being described by this use case scenario.

Figure 3.17 and the subsequent paragraphs describe the different interactions of the following actors.

- AGVs
- Surveillance Camera / Obstacle Detection System
- Other Actors

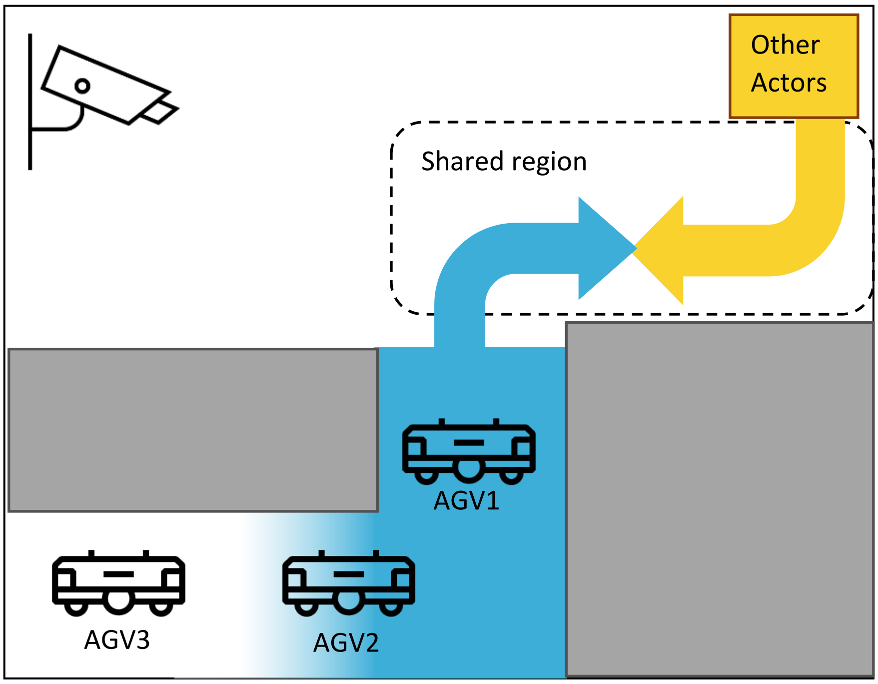


Figure 3.17: AGV obstacle detection

The application here showcases flexible groups of multiple AGVs (AGV1, AGV2, AGV3, …). Each AGV is used to transport various goods or parts. The parts themselves are not relevant for this scenario. These AGVs come from a hallway that obstructs the direct view / access to a shared region. This shared region can be accessed by the AGVs and other actors that might not be directly integrated into the higher-level factory automation system. Examples for such actors might be human personnel, human-controlled vehicles, etc. An obstacle detection system provides information regarding this shared region and potential objects within it. This obstacle detection system might use various sources of data that can be combined. This use case scenario uses a surveillance camera as example. Further sources may include a combination of multiple cameras, or additional sensing information provided by network participants.

The challenge in this scenario is that every AGV requires information provided by the obstacle detection system. As soon as an AGV approaches the shared region, it enters the so-called interest zone (marked as blue region in Figure 3.17). Within this interest zone, the need to have access to the obstacle detection information increases gradually as the AGV comes closer to the shared region. After the AGV passed through the shared region, the need for the obstacle detection information falls off and becomes obsolete.

Any AGV queued in the hallway, after the first one, can also benefit from the obstacle detection information. The idea is that the movement speed of an AGV could be increased if it is aware that no other actors will be present in the shared region. This flexible definition of the need for the communication required for accessing the obstacle detection information represent a possibility to find different levels of operation.

For example, if the wireless communication resources allow all AGVs to access the obstacle detection information, the movement speed of all AGVs could be increased and the resulting throughput is optimized. This setting of resource allocation can be further presented as operation mode that enables the highest level of operation for the AGV transport system.

With reduced availability of network resources, the number of AGVs that are granted access to the obstacle detection information in advance can be reduced as well. This would limit the number of AGVs that can increase their movement speeds and therefore decrease the overall throughput.

The minimal viable mode of operation would be that only the first AGV, closest to the shared region, is granted access to the information whether other actors are present in the shared region. This would limit all other AGVs to move with the lowest speed possible and would therefore result in the lowest throughput of the AGV transport system.

These different modes could be represented as different levels of operation. This type of flexible communication requirements could further be used to optimize the overall factory performance depending on available network resources. The potential impact on the factory performance, the preference towards the higher operation mode level, and the priority of the operation mode can be seen as additional parameter for a communication request.

Some examples for these different operation modes can be found in Table 3.6 as multiple different variants of the same obstacle detection communication.

UCS3: Mobile Processing Modules

This use case scenario shows the potential of a different kind of mobile device in a factory automation application. In contrast to the previous use case scenarios, where the mobile modules were Autonomous Guided Vehicles (AGVs), this use cases scenario shows Mobile Processing Modules (MPM).

MPMs can be seen as AGVs that include additional technical functionality, like a robotic arm or other tools. Figure 3.18 depicts such MPMs, together with a stationary production line and an MPM base station as additional actors. The main motivations for using such MPMs are the enabling of shorter down-/changeover times and the elimination of human interactions when e.g., adapting the configuration of a production line for different batch sizes or product customization.

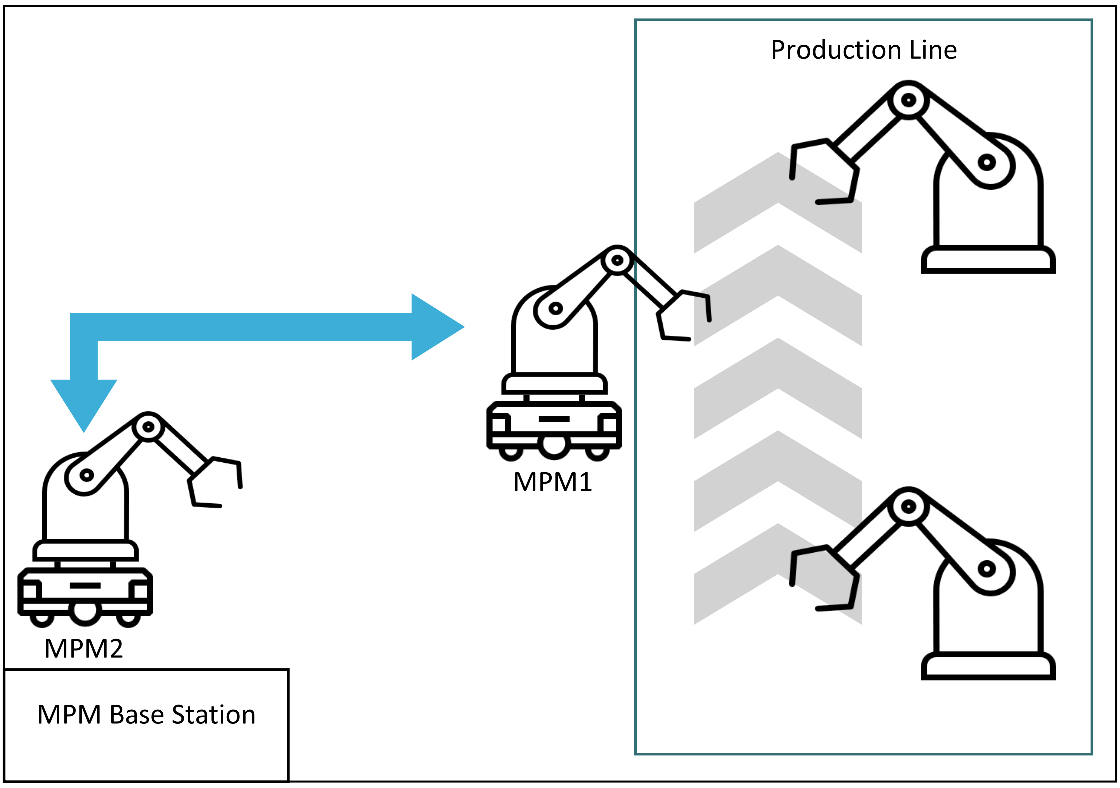


Figure 3.18: Mobile Processing Modules

MPM1 participates in the processing operation of the production line. During this cooperative processing a reliable communication between the production line and MPM1 is required. This communication allows MPM1 to be fully integrated into the processing operation. For the duration of this processing operation this communication remains unchanged. The physical movement of MPM1 is also either strongly limited or even fixed during the processing operation. This limited movement also has the potential to reduce some variable aspects of the wireless communication and the achievable properties. One example is the expected variation of the packet transmission. This may be lower in comparison to a communication in between fast-moving device. Therefore, this property of expected physical movement is also relevant for the communication requirements.

After a given processing operation is completed and a change of operation is being performed, MPM1 can logically detach itself from the production line. MPM1 could be replaced by (a different) MPM2, or the additional MPM2 also joins the production line for the next processing operation. Changes like this can have many different reasons, e.g.:

- Replacement of tools due to deterioration
- Cleaning and maintenance of included tools
- Change of tools due to different requirements for different products
- Change of processing operation that requires a different set of MPMs

During the time an MPM is not participating in the processing operation, the MPMs may interact with each other, like the AGVs described in the previous use case scenarios.

The MPM base station and its interactions with MPM2 can be viewed similarly to the production line and MPM1, although the criticality of the required communication with the MPM is expected to be lower. The communication between MPM2 and the MPM base station may serve the following purposes:

- Exchange of status information about the MPM
  - Battery level
  - Estimated or measured state of mounted tools
- Exchange of control information regarding the functionality of the MPM base station
  - Control of required operations like tool change and maintenance
  - Control of charging operations

The following paragraphs summarize the different communication relations for each mode of operation. More detailed information and example values for different communication properties can be found in Table 3.6.

MPM docked to the production line:

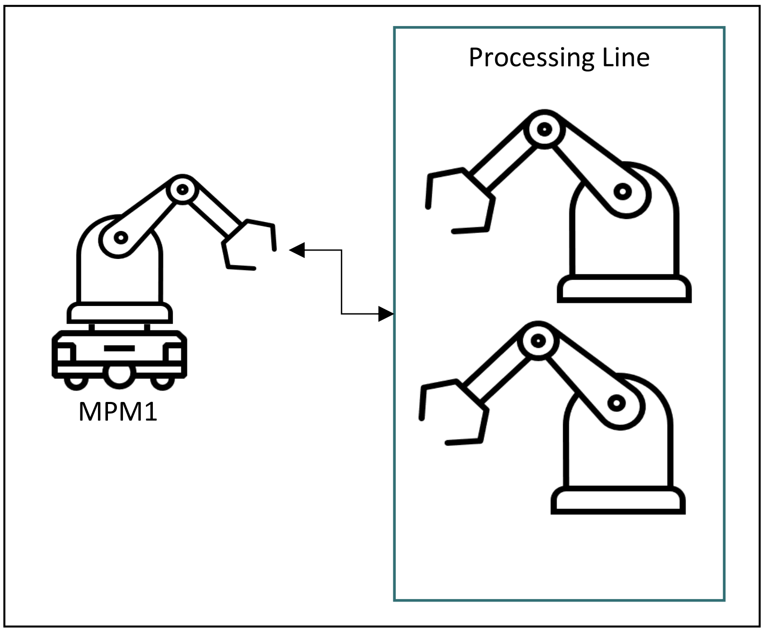


Figure 3.19: Communication between MPM and production line

The registration / deregistration via aperiodic communication matches other registration / deregistration communication relations shown in previous sections and use case scenarios.

High priority communication is required between MPM1 and the production line during combined processing operation. This communication requires strict timing guarantees to fully integrate the MPM into the production line. This communication usually follows a periodic communication scheme, where the reliability is essential for fast coordinated operations.

MPM moving next to other MPMs:

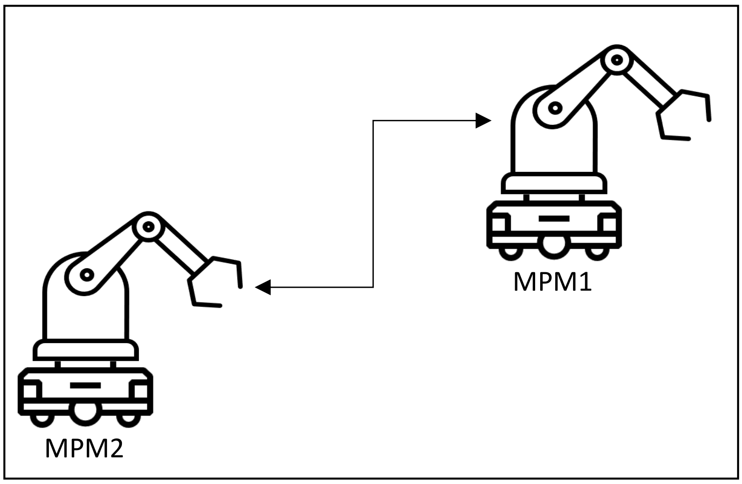


Figure 3.20: Communication between MPMs

The communication required during the movement phase of the MPM is like the previous coordinated movement communications, with the only difference being that this communication is less time critical as the different MPMs do not move as a coordinated group, but as different individuals that inform each other about their position and planned movement. This type of information could also be delivered via an aperiodic communication.

MPM docked to MPM base station:

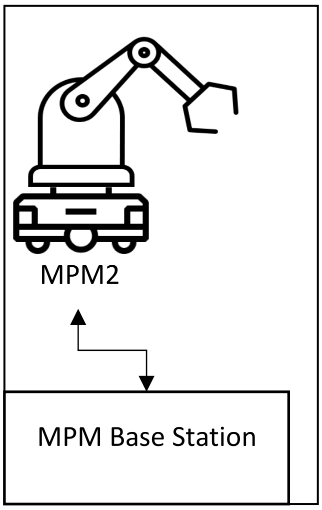


Figure 3.21: Communication between MPM and base station

The operation mode of an MPM being docked to its base station and the different types of required communication relations resemble many previously shown communication relations. This includes an aperiodic communication for registration and deregistration.

Other communication could include the exchange of status and/or control information. The required communication properties strongly depend on the implemented logic. Usually, these types of information exchange could be implemented using either aperiodic or periodic communication. For all types of communication between the MPM and its base station a rather low priority is sufficient as the performed operations resemble background tasks and are not as critical for the overall factory operation as other communications.

UCS4: Combination of collaborative robots with Extended Reality and Exoskeletons

This use case scenario is an addition to the previous scenarios, which not only focuses on applications within the field of factory automation, but also shows different possibilities to combine the fields of Factory Automation, Extended Reality, and Exoskeleton. More information about specific use cases and use case scenarios related to Extended Reality or Exoskeleton can be found in Section 3.2 and 3.3.

The main starting point of this scenario and potential future applications is the usage of a collaborative robot, like the ABB YuMi[3] in combined operation with a human worker. This cooperation requires bidirectional communication between the human worker and the collaborative robot. This type of interaction can be vastly improved by using technologies like Extended Reality or Exoskeletons. The combination with each technology already exists today, but the actual applications are often limited by either the need for wired communication or by limitations of the currently available wireless communication technologies like Wi-Fi.

shows an exemplary scenario where two human workers cooperate with a collaborative robot to perform an operation on a part in the middle. Each worker wears either an HMD or an Exoskeleton. The simultaneous usage of both devices by a single worker is also imaginable, but the depiction of this combination was omitted for simplicity. This whole application, including all different actors, is also being represented within a digital twin. The depicted connections represent bidirectional communication relations between the different actors.

[3] https://new.abb.com/products/robotics/robots/collaborative-robots/yumi/irb-14000-yumi

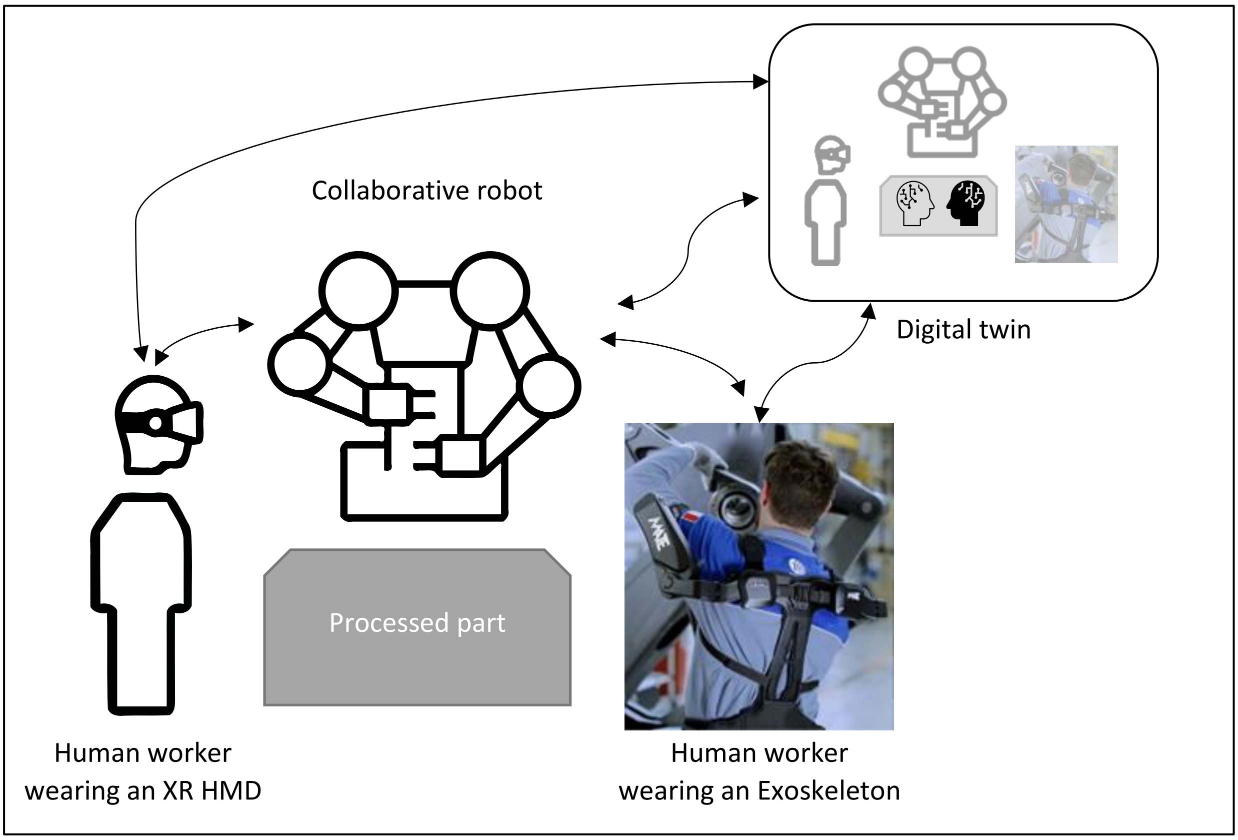


Figure 3.22: Interactions within the combined use case scenario

The following paragraphs describe some potential future applications to show what kind of new innovations could be enabled by the DETERMINISTIC6G project. These can be visualized as different parts, or scenarios of the application in Figure 3.22.

The combination of a collaborative robot, or any other type of collaborative machine, with an Exoskeleton could be done for different purposes improving the collaboration or allowing different types of interactions. For example, the Exoskeleton could be used as an actor that protects the human worker. This protection could be done via a guided support that leads the human worker away from potentially dangerous scenarios. Such a guided support could for example be implemented by stronger physical support by the Exoskeleton for safe movements and weaker support, or even increased resistance for unsafe movements. Such unsafe movement could be an approach towards a safety-critical zone containing threats like moving parts, high temperatures, or high voltages.

Another use of the Exoskeleton would be the usage as sensor or generalized as source of information. This information could include the current position and movement of the human worker. This information about the current and the intended movement could furthermore be used to control the collaborative robot to correctly react to the movement and for example correctly move or manipulate a part that is held by both the collaborative robot and the human worker.

The integration of an Exoskeleton as both actor and sensor could enable various future applications. Another example would be an estimated location of the human worker, not only limited to the geographical location within a factory, but also the location, orientation, and reach of different extremities. This could be done via a simulation of a digital twin of the human worker that is fed with information from different sources like the Exoskeleton itself and maybe additional camera data. The use for such a simulated model could be again safety-relevant checks that protect the human worker from approaching dangerous zones or performing dangerous operations.

Similar to the potential integration of an Exoskeleton some types of Extended Reality devices could also be used as sensors or monitors. A Head Mounted Device (HMD), as described in the previous use case Section 3.2 could be used to display relevant information. There are many different uses for such a direct display, for example displaying the intended movement, or operation of the collaborative robot. This could benefit the human worker and allow a better coordination with the collaborative robot.

Another future-oriented idea is the incorporation of prediction data regarding the expected communication properties into the application itself. This prediction data could come from some model that predicts the available communication resources and the achievable guarantees for a given time frame in the near future. This prediction data could be used to prepare any ongoing operation like a coordinated movement of a human worker and a collaborative robot. Such a reaction could be for example a warning about an expected degradation of the communication level and maybe a reduction in operation and movement speed. Such warnings could again be transported either via physical feedback of the Exoskeleton, or via an XR display.

These and many more usages and potential applications could be envisioned using the technologies of XR, Exoskeleton, collaborative robots, and reliable wireless communication. This list of examples shall only provide a creative starting point for future applications.

3.4.4 Communication relations and requirements

Table 3.6 shows the summarized communication relations from the previously shown use case scenarios with some representative values for different KPIs. These values shall visualize the potential ranges for different communication relations. The specific values are not related to an actual application as the variability between different applications is too wide.

The place holders used in this table of communication relations shall be interpreted as follows:

- n: generic number of devices
- N: significantly higher number of devices than n
- X: single digit value
- XX: double digit value

Table 3.6: Communication relations with KPIs for the Factory Automation: Adaptive Manufacturing Use Case

Communication content	Communication purpose	Recurrence	Period (ms)	Bandwidth (MBits/s)	Communication cardinality	Max Transmission Latency/ Packet Delay (ms)	Min Transmission Latency/ Packet Delay (ms)	Transmission Jitter/ Packet Delay Variation (ms)	Max Packet Loss per unit of time	Prioritization (of Communication Relationships) / Ranking	Communication Duration	Distance/ Position/ Velocity
Status and Control Information	AGV Movement Coordination	Periodic	10	medium	1:1	10	-	5	2 / 30 ms	1 / High Priority	Continuous	0 – 2.2 m/s
Control Information	Functional Safety Stop	Periodic	20	low	1:1	10	-	10	2 / 60 ms	0 / Highest Priority	Duration of movement through safety relevant zone 10s – X min	0 – 2.2 m/s
Device Information	Service Registration / Deregistration	Aperiodic	-	low	1:1	>100	-	-	-	2 / Medium Priority	Short burst	0 – 2.2 m/s
Status and Control Information	AGV / Processing Line Movement Coordination	Periodic	5	medium	n:1 1:n	2.5	-	1	2 / 15 ms	0 / Highest Priority	Depends on performed operations 10s – X min	0 – 2.2 m/s
Obstacle Information	Minimum Level of Obstacle Detection	Periodic	10	low	1:1	5	-	1	2 / 30 ms	1 / High Priority	Duration of presence within the interest zone	0 – 2.2 m/s
Obstacle Information	Medium Level of Obstacle Detection	Periodic	10	medium	1:n	5	-	1	2 / 30 ms	2 / Medium Priority	Duration of presence within the interest zone	0 – 2.2 m/s
Obstacle Information	High Level of Obstacle Detection	Periodic	10	high	1:N	5	-	1	2 / 30 ms	3 / Low Priority	Duration of presence within the interest zone	0 – 2.2 m/s
Status and Control Information	Combined Processing Operation	Periodic	1	medium	1:1	1	-	0.5	2 / 3 ms	0 / Highest Priority	Duration of combined processing 10s – XX min	0 m/s, still standing

3.4.5 Novel challenges

Current technologies in the field of Factory Automation often rely on statically allocated communication resources. This strategy is often suitable for wired communications and fixed applications, but as soon as more flexibility is required by an application, the traditional strategies for resource management become inefficient or unusable. This need for flexibility is emphasized by using open and more broadly accessible communication technologies like 6G. At the same time, there is a high demand for low communication periods and strict timing.

In the context of DETERMINISTIC6G, the main requirements for wireless communication in the use case Factory Automation: Adaptive Manufacturing are:

- Low latency for wireless communication
- Dynamic reallocation of critical communication resources
- Coexistence of different types of communication without negatively affecting high priority communication

These requirements can be used to identify potential solutions or strategies for finding solutions by relating them to the key ingredients mentioned in Section 1.1. The ingredients (1) – Defined requirements, computational footprints and level of reliability - and (3) – Adaptive resource allocation and dynamic applications - are foundational functions to enable a dynamic reallocation of critical communication resources. Ingredient (2) – confidence to support E2E deterministic communication - represents the base for achieving a low latency.

3.4.6 Security requirements

Factory automation networks are usually detached from the public internet, and remote interactions are often added via separate gateways. Nevertheless, the tendency is towards more open systems, the use of wireless communications (e.g., IoT and 5G networks), and the adoption of virtualisation technologies (e.g., cloud computing and multi-access edge computing (MEC)) that introduce new vulnerabilities.

The security requirements depend on the scenarios and applications and can vary. In some cases, maliciously inserted flawed information could lead to various errors and failures that have the potential to destroy used materials or the machines themself. A use case that can be considered critical is, for instance, Automated Guided Vehicle (AGV) were the manipulation of the information processed to detect obstacles (e.g., the camera video stream) could provoke accidents or impact the services. Another use case involves Unified Control Systems were the manipulation of the control information communicated to each entity could lead to erroneous operations, faulty products, or more serious accidents involving humans and machinery.

As indicated before in other use cases, time synchronisation can be important, but this depends on the operation mode. For instance, an MPM docked to the production line might require more accurate synchronization than an MPM moving back to the base station, or an MPM that resides at the MPM base station.

Depending on the type of use case, packet loss, packet delay variation, and synchronization accuracy, the communication may need to be monitored to detect any deviations from expected behaviour. The expected behaviour can usually be derived from the implementations of the application plus the configuration of the system.

The targeted automation of the manufacturing processes requires that the management of security (i.e., prevention, mitigation, response to security breaches) can be automated as much as possible. This could also include human confirmations or interventions.

With the emerging shift towards the Industrial Internet of Things (IIoT) and increased connectivity, applications cannot rely on the currently common physically isolated network architecture anymore and have to consider security aspects. This requirement for security gets amplified by using more open and farther-reaching communication technologies like wireless communication and especially mobile networks. Creating physically separated networks using 5G or 6G technologies is often either not possible or simply economically not feasible. Therefore, other strategies gain importance to ensure the correct and reliable operation of a factory. Some of these strategies are listed below.

- Logical / virtual separation of shared communication resources
- Encryption of transmitted data
- Validation / verification of sender and receiver roles

3.5 Mobile Automation: Smart Farming

3.5.1 Overview

Agriculture is the foundation of our society. However, supplying the world's population with sustainable food sources will be a huge challenge. It is estimated that the earth’s population will be 9.7 billion by the year 2050 and as a result, the level of automation must increase proportionally. In addition, this must be done while preserving the environment.

In the future, climate change will give rise to unexpected weather conditions which will entail severe food shortages. The greenhouse gas emission is raising global temperature, increasing pest and weed infestation and altering precipitation patterns. All these challenges can lead to a change in the type of crops grown and the time of the year they are harvested. Moreover, as soil is becoming less fertile and eroded, it will be more difficult to increase or even maintain yield. Additionally, increasing urbanization will mean less land will be available for farming. This, in combination with the aging workforce in some regions, will add increasing difficulties in ensuring an uninterrupted global food supply chain. This implies that efficient agricultural farming operations and water resource management are required so that output from existing land is maintained or increased by reducing resource consumption, such as water, and taking preventive measures to protect the crops from infestation.

Farming today globally contributes to the green gas emission by 10% and ways to reduce this footprint will be necessary. One method to achieve this is to automate and optimize the complete farming process, as well as replacing the use of fossil fuels with electrification. The forthcoming challenges related to food production can be partially addressed by using automation-based smart farming techniques, which will enable efficient farming operations in terms of quick identification and treatment of damaged crops, autonomous farming vehicles, which communicate with each other in farming operations such as harvesting, ploughing, sowing, or spraying, availability of real-time information from the field to be used to update the planning of farming activities, remote control of farming vehicles, etc. Smart farming refers to the automation of farming activities and the processing of data for diagnostics and treatment of fields, automating recurring and hazardous tasks, predicting potential problems and maintenance issues, as well as implementing techniques to keep crops safe from unpredictable weather and pests.

This will reduce waste and help in increasing yield, without increasing required field area. Automation-based farming utilizing IoT, 6G, AI, and edge-computing based technologies can significantly improve the overall farming process. Some of the automation-based farming processes will require reliable and deterministic communication between involved entities, which may be enabled by technologies such as 6G.

3.5.2 State of the art

Automation-based farming or smart farming will to a large extent depend on autonomous farming vehicles such as harvesters, plows, rollers, planters, sprayers, etc. These autonomous vehicles are expected to be fully autonomous in their operations and will also be expected to collaborate to perform a task. There are already companies[4] producing fully autonomous tractors that enable the operator to process a field in straight and uniform parallel rows for sowing and other activities with an accuracy of 5 cm. These vehicles[5] use satellite communication to achieve the required accuracy. Similarly, robotic components, such as vehicle control units[6], already exist which can be installed into any vehicle platform, converting it from manual to automatic/autonomous control. These vehicles can be used in different applications, e.g., mining, farming, automotive, etc.

Light sport aircraft, drones, and autonomous robots can provide high-resolution images of the complete field, down to leaf level. Some of them provide high-resolution images at speeds of 160 km/h, rapidly identifying, e.g., zones where there is infestation. By running high granular field-level weather forecasts and using machine vision and machine learning to collect and interpret field data into useful information, it is even possible to recommend the most effective time to apply treatments or the best planting window[7] [Res20].

To support sustainable smart farming, accurate energy management is required, where an energy manager monitors the complete spectrum of energy consumption and electrical distribution assets on the farm, including HVAC and pumps, to maintain efficiency and availability for a healthy harvest [Sha23]. There also exists a wide range of wireless-based mesh routers, gateways, sensors, meters[8] etc., which can be used in agricultural applications[9]. Today, there are several companies offering products that can be essential for smart farming, such as autonomous farming vehicles, high-resolution imagery using drones, autonomous robots to collect field information, machine learning and AI algorithms to extract useful information from the collected data, sensors, gateways, etc.

[4] https://www.deere.com/en/index.html
[5] https://asirobots.com/farming/
[6] https://www.br-automation.com/en-in/products/plc-systems/x90-mobile-control-system/
[7] https://www.taranis.com/acquisition/
[8] https://www.wittra.io/products/wittra-mesh-router/
[9] https://www.wittra.io/use-cases/agriculture/

Autonomous machines are also already used in other domains like mining and the forest industry. In underground mining, operations like loading and transporting ore are done autonomously. Self-driving machines that pick-up timber can be used by wood-processing companies[10].

Unmanned Aerial Vehicles (UAVs) are widely used across different industries for tasks like inspection, light transportation, or mapping. While the basic components of smart farming are already in place, a completely autonomous farm will require much more automation and coordination between different entities such as Unmanned Ground Vehicles (UGVs) and UAVs.

Many of these components represent quite advanced technology and it is critical that the complexity of these systems is hidden and that it is possible to operate them by non-experts, since a requirement for highly trained professionals would make the employment of such technology too costly.

A secure and robust communication infrastructure must be in place, which can provide deterministic, high-bandwidth, and long-range communication for a very diverse set of applications. Upcoming communication technologies such as 6G, with great potential to be more deterministic, can be enablers for smart farming applications, like outlined in the following sections, but also for many other application areas in the mobile automation domain. Edge computing will provide even more opportunities to offload computationally expensive algorithms from the vehicles or AVs to reduce power consumption on autonomous vehicles in the field.

3.5.3 Use case description

The following smart farming use case provides an abstract example of the possibilities and trends in this specific application area. Similar usage scenarios can be found and realized in other mobile automation domains, like digital construction, mining, or embarkation. Figure 3.23 shows the main entities and the communication scenarios of an exemplary autonomous and sustainable smart farm, which also has an option for remote operations. The UGVs, e.g., harvesters, rollers, planters, etc., are envisioned to be autonomous or remote-controlled. There are UAVs such as drones or light sport aircrafts, which can also be autonomous or remote-controlled. It is envisioned that real-time communication is required among the different UAVs and the UGVs to collaborate on a farming operation. Moreover, real-time information is required to be exchanged between the vehicles, and between the vehicles and a remote-control application to avoid collision with field personnel, other vehicles, animals, or obstacles of any kind, and to optimize operations. To offload computationally intensive applications from the vehicles, for instance to reduce power consumption and extend the duration of operation, edge computation is the proposed solution. UAVs, UGVs, robots, and sensors can collect data from the field that may later be used to identify a problem area such as pest infestation and weeds or to identify areas with need for irrigation or fertilization. The exemplified farm also includes a control centre, where an operator can monitor operations if required. It is also expected that corresponding management applications distribute updated processing plans, maps, or software updates to the UGVs and UAVs.

Self-sustainability will be desired for remote farms, especially those which are not grid-connected, and which are dependent on their own power generation. Integration of renewables (wind farm or solar panels) can reduce dependency on fossil fuels. However, variable power production from renewables

[10] https://www.einride.tech/press/all-electric-autonomous-logging-truck-revealed-at-goodwood-festival-of-speed/

requires energy storage for stable grid operation. UGVs may run on electricity and in some cases act themselves as energy storage or there may be additional storage facilities at the farm. In this case, a power management system is required to balance demand and supply of electrical power in the farm. In some cases, the farm's electrical system may be grid-connected. Then, the occurrence of malfunctions in the system may require the farm to be electrically disconnected from the grid. This requires fast information exchange between electrical relays.

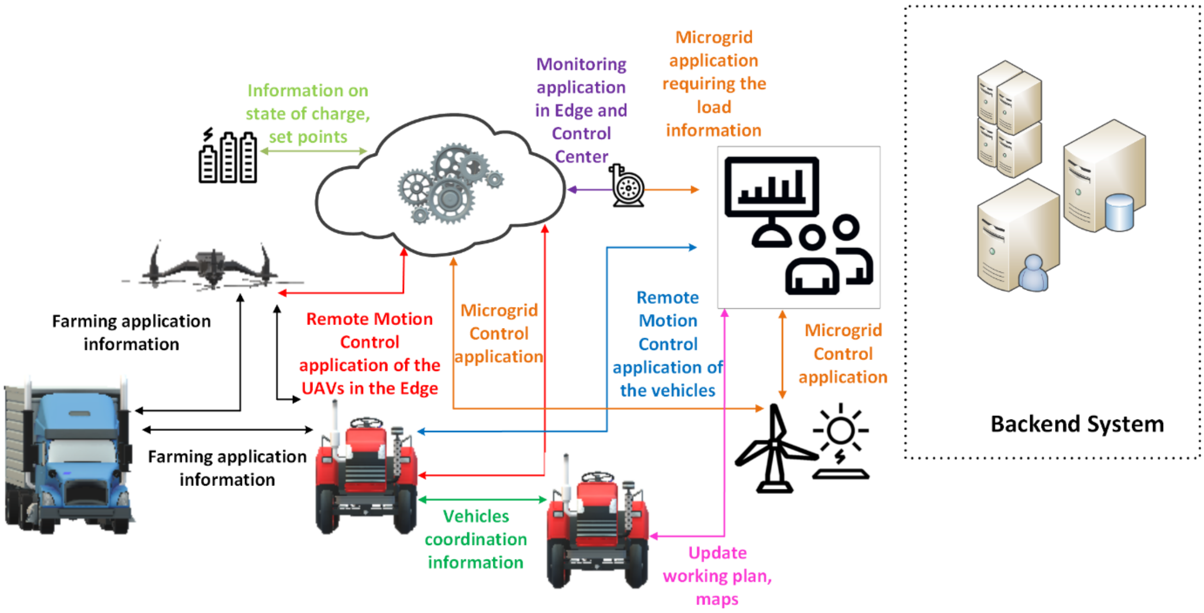


Figure 3.23: Smart Farming communication scenarios

Figure 3.24 shows an application executing on different entities such as the Edge or a control centre. Applications which are computationally intensive can be offloaded to the Edge. For instance, applications such as those used for path planning would be used to pre-determine the navigation path of GVs and AVs. Similarly, based on collected data from sensors, AVs, robots, etc., monitoring applications may predict possible problem areas and take countermeasures to prevent damage and waste. Offloading safety-critical applications such as collision-detection from the AVs to the Edge can reduce power consumption and extend battery life of AVs, such that they can be used for an extended duration.

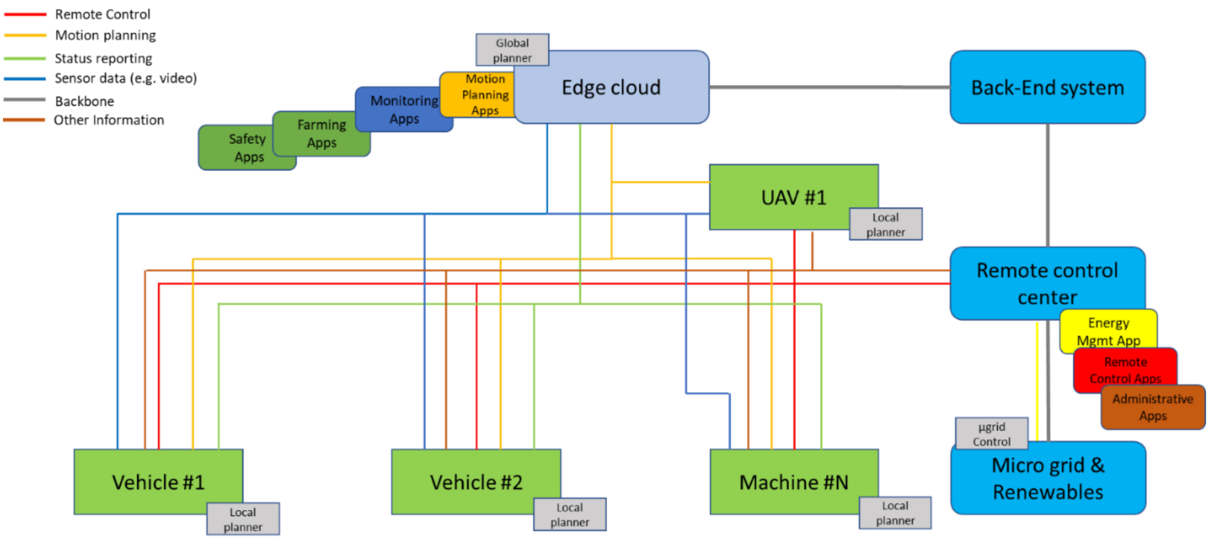


Figure 3.24: Smart farming applications

UCS1: Combine harvester farming operations

A combine harvester is a device that cuts and delivers the grain or seed crop to a threshing machine, which separates the seeds from the crops and works as it moves across the field. The yield from the harvester can be emptied into a trolley as it moves across the field. Today, vehicles are driven by human operators, but in the future, they are expected to be autonomous. Thus, close coordination and communication is required between the harvester and the trolley into which the yield is emptied so that no yield is wasted, and continuous operation is maintained. It is expected that communication between vehicles, such as the harvester and the trolley, is time-synchronized and location information is retrieved from either cellular networks or satellites.

The following information is required to be exchanged between the harvester and the trolley as shown in Figure 3.25 and summarized in Table 3.7.

1. Handshake data: Before the harvester starts emptying the yield into the trolley, the following data is required to be exchanged: mutual authentication, estimated time left to start offloading of the harvest, timestamped coordinates, speed, and acceleration of the vehicles. If the harvester is already offloading to another trolley, then the trolleys need to mutually authenticate each other and exchange information such as timestamped coordinates, speed, and acceleration with each other to avoid collision and maintain coordinated takeover.
2. Coordinated movement data: Once the initial handshake has been performed and offloading starts, the harvester and the trolley keep exchanging status information such as timestamped coordinates, speed, and acceleration with each other to maintain proper distance. If there are consecutive trolleys to take over offloading, then similar data is exchanged between the consecutive trolleys. Other than this, start/stop commands, signalling the start and stop of the offloading, are sent to the trolley.
3. Closed-loop control data: Information such as amount of empty space left in the trolley is periodically sent from the trolley to the harvester. The frequency of this data exchange is reduced when the remaining empty space reaches a certain threshold. Depending on the amount of remaining space, the harvester can reduce the flow of yield to the trolleys.

4. Connect/Disconnect command data: Once a trolley is full, it disconnects from the harvester and other trolleys and moves to another location for further processing of the harvest. In this case, multiple trolleys may subscribe to a signal from the harvester which indicates that the currently active trolley is about to be full and shortly disconnecting. Upon reception, an idle trolley may signal its availability and take over.

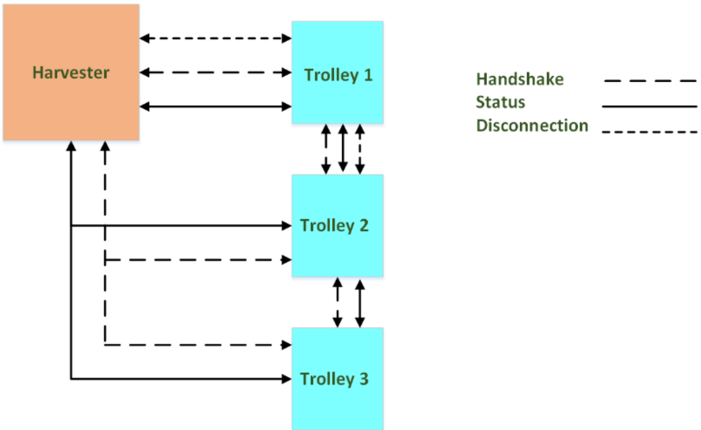


Figure 3.25: Information exchange in combine harvester farming operations

This use case demonstrates the need for a communication network to support real-time closed-loop control farming applications. The status of vehicles such as speed, acceleration and time-stamped coordinates can be exchanged through publish/subscribe mechanisms. If for some reasons the vehicles do not receive status information, they are expected to reduce speed or come to a complete stop to avoid accidents. Highest priority should be given to the requirements of this communication relation to prevent interruption of farming activities and ensure safety.

Table 3.7: Information exchange in UCS1

Type of data	Information	Actors Involved
Handshake	Mutual Authentication	Between trolleys
		Harvester and all the trolleys
	Estimated time left to start of offloading of the harvest to the next trolley	Harvester and the next trolley
	Timestamped coordinates, speed, and acceleration	Harvester and trolleys
		Between the two consecutive trolleys
Status	Timestamped coordinates, speed, and acceleration	Between two consecutive trolleys
		Between the harvester and the trolleys
Closed-loop Control	Signalling the start and stop of the offloading to the trolley	Between the harvester and the current trolley
	Amount of remaining empty space in current trolley	Between the harvester and the current trolley
Connect/Disconnect commands	A signal is sent to inform that the current trolley connects or disconnects	Between current trolley to the harvester and the next trolley

UCS2: UAV-supported harvesting

While in UCS1 the UGVs are exchanging status information with each other to enable them to maintain coordinated movement, there may be obstacles on the pre-planned path which would require the UGVs to slow down or stop. Examples include crops of different size, animals, or other vehicles with which the UGVs do not have pre-established communication. In the future, it is envisioned that UGVs will be equipped with technologies such as LIDAR which will enable them to determine the distance to obstacles. However, the type of obstacles will be difficult to determine with this technology. Classifying the type of obstacles will be important to take appropriate measures. For example, detecting obstacles which are moving at moderate speed, e.g., other farming vehicles, will impose higher requirements on communication, i.e., low-bounded latency, than the detection of small static obstacles on the path. In smart farming, UAVs are envisioned to be used to support different farming applications, e.g., VRT (Variable Rate Treatment, applies fertilizers, seeds or pesticides in a variable rate, as opposed to uniform application). Again, for safety-critical applications, it is assumed that the UGVs and the UAVs are time synchronized.

Figure 3.26 shows the information exchanged between UAVs and UGVs, such as the harvester or the trolley. On the one hand, the UAV takes images of different resolutions, e.g., for safety-critical applications, medium-resolution images may be taken by the UAVs for obstacle detection. The images may be transmitted periodically to the harvester and the trolley. AI algorithms determine the type of object and act accordingly. On the other hand, high-resolution images can be transmitted, showing leaf-level details which may indicate required treatment of infestation and decay applying VRT. Other than this, information related to the status (speed, acceleration, and time-stamped coordinates) is expected to be exchanged periodically between UAVs and UGVs to maintain coordinated movement.

Safety-critical or farming applications based on image processing from the UAVs may be computationally intensive. In this case, image processing can be offloaded from the UGVs to the Edge. In zones with limited signal reception or high interference caused by environmental factors, available bandwidth may be reduced. In such cases, UAVs may degrade the quality of transmitted images (e.g., reduce image resolution), which may cause vehicles to move slower, but still maintain safety guarantees.

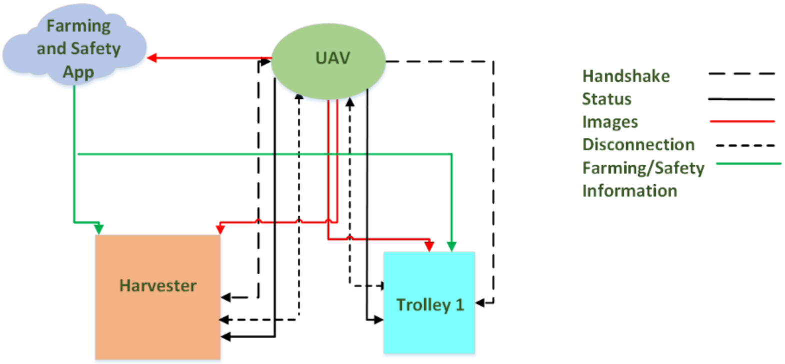


Figure 3.26: Information exchange in UCS2

UCS3: UAV and UGV based remote control

UAVs may produce high-resolution images or video streams, which can be used in many different applications, e.g., VRT. Similarly, UGVs such as robots can be used to collect information which can be used in monitoring applications. In some cases, an operator may be directly involved by remote controlling the UAVs and UGVs. For example, an operator may control vehicles in the field or the UAVs and manually interact with video/image hardware to retrieve required information, such as soil condition or active vehicles on the field. Continued low-bounded latency communication is required for remote controlling to support fast response times of moving vehicles in potentially hazardous weather conditions or to avoid damaging crops.

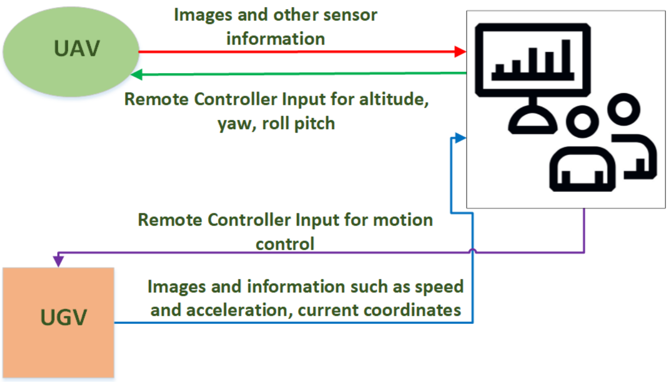


Figure 3.27: Information Exchange in UCS3

3.5.4 Communication relations and requirements

Table 3.8 shows the communication relations and their properties. Actual values may be influenced by environmental factors or be subject to the concrete use case.

Table 3.8: Communication relations with KPIs for the Mobile Automation: Smart Farming use case

Communication content	Communication purpose	Recurrence	Period (ms)	Bandwidth (MBits/s)	Communication cardinality	Max Transmission Latency/ Packet Delay (ms)	Min Transmission Latency/ Packet Delay (ms)	Transmission Jitter/ Packet Delay Variation (ms)	Max Packet Loss per unit of time	Prioritization (of Communication Relationships) / Ranking	Communication Duration	Distance/ Position/ Velocity
Video/Images	Identifying obstacles on the path	Periodic	2-100 Inversely proportional to the speed of UGVs	High	1:N	Typically, 50% of the period is worst case	Lower latency improves obstacle detection	Lower jitter improves obstacle detection	0	High Priority	Continuous (until a replacement UAV arrives, or operation is finished)	Few meters to 600m 0-80 km/h (UAV) 0-50 km/h (UGV)
Vehicle status information	To maintain coordinated movement between vehicles	Periodic	2-20 Inversely proportional to the speed of UGVs	low	N:N	Typically, 50% of the period is worst case	Lower latency improves vehicle safety	Lower jitter improves vehicle safety	0	Highest Priority	Continuous	0-50 km/h (UGV)
Application status information	To indicate amount of remaining space in current trolley	Periodic	50-100	low	1:1	100	20	10	Application can tolerate 1 to 2 consecutive packet losses	Medium Priority	Continuous	5m-10m between the harvester and the trolley 0-50 km/h (UGV)
Connect/Disconnect command	A command is sent to the harvester to indicate that the current trolley will move away	Aperiodic	-	low	1:2 Between current trolley to the harvester and the next trolley	100	-	10	Application can tolerate 1 to 2 consecutive packet losses	Medium Priority	Short burst	5m-10m between the harvester and the trolley 0-50 km/h (UGV)

Obstacle notification	Data from the Edge to the vehicle(s) about obstacles on the path	Aperiodic	-	low	1:N	Depends on the distance and speed of the vehicles. Low latency communication is desired	Lower latency improves reaction time on detected obstacles	Lower jitter improves performance	0	High Priority	Short burst	0-50 km/h (UGV)
Remote control data	Remote control input for motion control	Periodic	2-20 Depends on the current conditions, e.g., weather, rough terrain, etc.	low	1:N	Typically, 50% of the period is worst case	Lower latency improves real-time motion control	Lower jitter improves real-time motion control	0	High Priority	Continuous	0-80 km/h (UAV) 0-50 km/h (UGV)
Monitoring data	Monitoring information from different sensors	Periodic	≥100	Medium	N:1	20	10	10	Can tolerate 4-5 consecutive packet losses	Low Priority	Continuous	0 m/s

3.5.5 Novel challenges

Current and upcoming cellular technologies, such as 5G, have started to move in the direction of providing deterministic and low-latency communication to enable machine-to-machine communication but still have not achieved the required performance for time-critical control. 6G has the potential to provide determinism and reduced packet delay variation, both of which is a requirement of many control applications in the mobile automation domain. In this domain, one of the challenges to be solved is to ensure the reliability of control loops, which are needed for collaboration, and which are closed via a wireless communication system. At the same time, flexible interaction of autonomous vehicles needs to be enabled.

Thus, smart farming applications will especially benefit from capabilities, like accurate time synchronization, bounded low latency and centimetre-level location information of upcoming cellular technologies such as 6G. These capabilities are also reflected in the key ingredients of the project goals described in Section 1.1, with the ingredient (2) - confidence to support E2E deterministic communication - providing most benefits for the illustrated smart farming use case. The goal is to implement these ingredients, such that in future supporting automated and autonomous operations will be possible to a larger degree than today.

3.5.6 Security requirements

Every wireless network is vulnerable to signal jamming. This is also true in case of cellular networks. More specifically, the uplink signal can be jammed more easily than the downlink signal [AA22]. Even though wireless communication technologies such as 5G or 6G will make use of technologies such as beamforming, which in some ways may mitigate jamming effects, the possibility of attacks remains. In other use cases, where communication takes part indoors such as communication in a factory building, wall penetration loss caused by buildings can effectively improve resilience against jammers located outside the factory. This is not the case in the smart farming use case where UGVs and UAVs operate out in the open.

Similarly, an attacker can transmit fake Primary Synchronization Signals at high power to conduct DoS attacks against legitimate entities during initial cell search. Other types of critical attacks include replay attacks, where the content of a communicated message is modified and sent to a legitimate receiver. This can be a problem when communicating information such as vehicle status and obstacle detection information which may be used in safety-critical applications.

Protecting the network from being compromised can be a challenge, especially since network virtualization may introduce additional vulnerabilities related to programmability of network functions. Examples include adding malicious logical nodes into the core network, and manipulation of forwarding logic, etc., which may cause time-critical data to be delayed. Similarly, problems can arise when the farm’s control network is connected to the internet. Examples of an attack include injection of malicious code to extract sensitive information.

4 Key Values and Key Societal Value Indicators

While KPIs play a key role in the technical design of the solution, the design paradigm for 6G technology and the enabled use cases is not only focused on performance capabilities but also on the added value these use cases bring for society and our planet. This strategic shift implies that technologies and solutions should be developed aiming at addressing societal challenges, pain-points, and needs, creating value for society. The UN Sustainable Development Goals (SDGs) and the European Green Deal provide a basis for an understanding on societal values. A technology is valuable for society if it enables key values (KVs). In [6GIA22] the societal KVs that 6G can address are presented along with a roadmap for identifying societal impact and human needs. Examples of societal KVs are environmental sustainability, societal sustainability, economical sustainability and innovation, democracy, cultural connection, simplified life, digital inclusion, etc. Furthermore, two methods are proposed for measuring the value of 6G solutions with the help of Key Societal Value Indicators (KVIs). These are shown in Table 4.1.

Table 4.1: Methods for evaluating KVIs from [6GIA22]

Assessment type/phase	**Lower TRLs (Technology Readiness level) (i.e., early in the technology development)**	**Higher TRLs (Technology Readiness level) (i.e., later in the technology development)**
Subjective assessment	Trials, experiments, interviews	Questionnaires, interviews, focus groups
Objective assessment	Assessment by subject matter experts	Measurements on deployed networks

4.1 Key Value Indicator analysis

The purpose of the KVI analysis is to be able to point at expected value benefits from technology use and to provide a basis for a value driven design of technology. The methodology suggested in [6GIA22] and illustrated in Figure 4.1consists of four steps: (1) Identifying relevant key values, (2) Identifying the KVIs, (3) Determining the enablers and blockers of usage – the KV enablers, (4) quantifying of KVIs with KPIs.

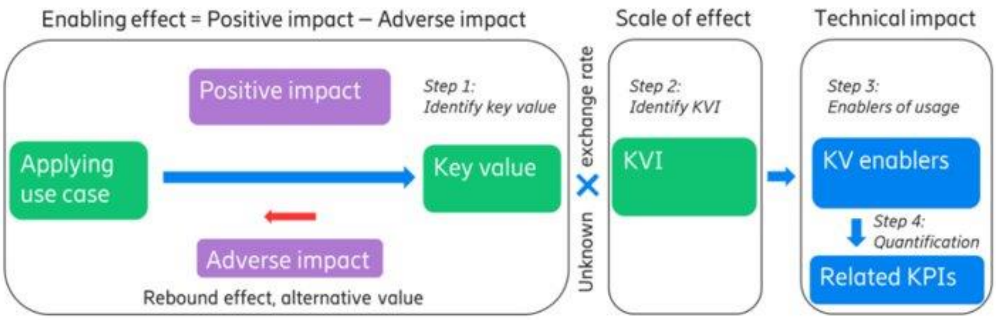


Figure 4.1: Overview of the KVI analysis in four steps [6GIA22]

The methodology taken in DETERMINISTIC6G builds on this framework but refines the impact area/sustainability dimension and has its starting point in the use cases.

It has to be noted that the analysis of key values and key value indicators is still very much in its infancy. DETERMINISTIC6G can contribute with input for the analysis of the use cases considered in the project but does not claim to solve the complex problem. In particular the measurement of key values is still an open problem when it comes to interpreting and comparing results.

4.2 KVI analysis methodology of DETERMINISTIC6G

Understanding the societal impact of the use cases investigated in DETERMINISTIC6G requires a detailed analysis of the use cases. The following methodology is applied in DETERMINISTIC6G. It originates from [LÖV23] where potential positive impact is called a gain and negative impacts are called losses. In order to ensure a broad view of which areas/parts of society solutions might impact, in a first step (1) important impact areas are listed and further detailed and exemplified. Thereafter, (2) starting from the use case all areas which are impacted by the use case are identified. (3) Hypothetical gains and losses enabled by the ICT solution are listed for each individual use case. This also includes potential gains and losses for unintended use or intended misuse of the solution. (4) Furthermore, an understanding whether negative consequences are mitigatable or not is developed. (5) Finally, key values are identified for measuring the impact. Hence compared to Figure 4.1, all these are steps to take within the left-most box.

In the following tables the impact areas and sustainability dimensions considered in DETERMINISTIC6G are detailed, examples are given for all of the impact areas.

Table 4.2: Impact areas, environmental sustainability

Environmental impact areas	Examples
Emissions to water, air, and soil	Emissions of greenhouse gases, nitrogen, phosphorus, chemicals, etc.
Freshwater	Use of freshwater and protection of water resources
Land use	Land footprint
Biodiversity (planet protection)	Protection of forests, cryosphere, etc. Protection of animals (over exploration, etc.)
Energy	Overall energy use
Material resources	Resource efficiency in consumption or production Amount of waste and loss Circularity Prolonging lifetime of products Upgrade of infrastructures or retrofit of industries Use of metals, plastics, etc.

Table 4.3: Impact areas, socio-economic sustainability

Socio-economic impact areas	Examples
Housing	Access to adequate, safe and affordable housing with basic services Increase the resilience of houses Warning systems for disasters
Connectivity	Access to affordable, high quality, reliable, sustainable, and resilient ICT

Transport	Access to safe, affordable accessible and sustainable transport systems for all Improved freight transports Number of road accidents
Energy	Access to affordable and clean energy Reliability of energy/electricity systems Upgrade of technology for supplying modern, sustainable energy systems
Water and sanitation	Access to safe and affordable drinking water Water quality, water-use efficiency
Food	Access to healthy food for all Food waste and losses
Education	Access to education for children Access to continued life-long education (regardless of gender, age, education level) ICT skills Promotion of life-long learning
Healthcare	Access to quality healthcare services Deaths related to hazardous chemicals, emissions to air, water, and soil Capacity for early warning, risk reduction of global health risks
Income and work	Unemployment, employment Prevention of physical safety issues Promotion of safe and secure working environments Facilitation of reskilling and upskilling possibilities Inclusion at workplaces (gender, age, disabilities, etc.) Promotion of decent work hours Promotion of inclusive and sustainable industrialization
Well-being and culture	Rights to practice one's culture Protection for cultural heritages/culture Involuntary loneliness
Peace and justice	Violence in society, violence and abuse of all forms
Political voice	Democracy, transparent institutions
Social equality and inclusion	Equal opportunities, payments Migration and mobility of people Social, economic and political inclusion of all Inclusion in process
Gender equality	Gender equality Equal opportunities, payments
Privacy and integrity	Balance between transparency, traceability, and privacy Integrity of data
Personal freedom and fair treatment	Freedom of expression Equal treatment in society Freedom to movement
Resilience	Resilience and adaptive capacity Early warning systems Knowledge and education on climate change

The identification of relevant key values as well as their assessment needs to be done by subject matter experts as well as sustainability experts. A crucial step is to determine possible indicators for a quantitative impact assessment. There exist various databases already today that can be used for assessment and evaluation on society-level. One important source of data is for instance the SDG tracker (Measuring progress towards the Sustainable Development Goals - SDG Tracker[11]) which presents data across all available indicators from the ”Our World in Data”[12] database, using official statistics from the UN and other international organization. It has to be emphasized that in many cases quantification cannot take place until deployment and if a quantification is made upfront uncertainties have to be considered throughout the process.

Understanding and quantifying the impact of a use case on society can be very difficult. One way of breaking down the problem, applicable in manufacturing, is to take a stepwise approach where the factory and their workers and/or the ecosystems surrounding the use case are handled as a ”sub-society”. In this way a factory or manufacturing enterprise can evaluate and understand the impact on their own members as a first step before going to a more complex analysis for understanding the impact on society as a whole. In DETERMINISTIC6G this step-wise approach is taken. In the following analysis key value indicators are therefore provided both on factory level and society level whenever feasible. The KVIs listed in the tables below are by no means a complete list but should only be seen as suggestions.

4.3 KVI analysis XR in manufacturing

In the following section, the KVI analysis for the use case of “XR for remote working, remote training and remote worker support” is performed. For a more detailed description of the use case see Section 3.2. Tables are provided for all impact areas where important potential gains or losses have been found. The impact areas with respect to Table 4.2 and Table 4.3 Table where no gains or losses have been found are left out.

4.3.1 Socio-economic sustainability dimension

Housing

Key impact areas:	Access to adequate, safe, and affordable housing with basic services
Gains	By remote working/training via XR, people can live further away from the factory – hence a larger area to live in would lead to more affordable housing. Less urbanization – more flourishing countryside.
Losses	-
Key value indicators (factory)	Assessment of affordable housing in the wider area surrounding the factory.
Key value indicators (society)	Average housing prices across the country

[11] https://sdg-tracker.org/
[12] https://ourworldindata.org/

Connectivity

Key impact areas:	Access to affordable, high quality, reliable, sustainable, and resilient ICT
Gains	XR remote working requires connectivity at more remote places, this will increase the demand for better connectivity at remote places and could therefore lead to generally better connectivity at remote places for all.
Losses	-
Key value indicators (factory)	Proportion of factory personnel with sufficient connectivity at home for remote working using XR (grouped by living area).
Key value indicators (society)	Proportion of households with access to good connectivity.

Transport

Key impact areas:	Reduced road accidents
Gains	XR will permit more remote working, potentially leading to less travel and commuting and thereby reducing the risk of having an accident.
Losses	-
Key value indicators (factory)	Number of reported road accidents by factory personnel.
Key value indicators (society)	Number of reported road accidents.

Education

Key impact areas:	Learning and education
Gains	Improved possibilities for up-skilling and re-skilling. Improved ICT skills. Improved possibilities to get affordable training from a trainer located remotely.
Losses	When the capability to be able to handle an XR device becomes a pre-requisite for being employed this could lead to a clear disadvantage for people with less favoured background not being able to afford an XR device on their own. For people not being able to work with XR devices (e.g., due to eye illnesses, headaches, etc.) this can create a disadvantage.
Key value indicators (factory)	Evaluation of effectiveness of education and training with XR. Subjective perception and evaluation of level of education and vocational training. Subjective perception and evaluation of upskilling and reskilling possibilities. Number of workers successfully undergoing up-skilling or re-skilling using XR.
Key value indicators (society)	Regional disparities of access to education and advanced job training. Regional disparities of access to developing XR handling skills.

Healthcare

Key impact areas:	Mental status, general health
Gains	With more individualized support via XR it is possible to meet humans on their individual development level when learning a new task, etc. which could lead to reduced stress levels. An XR overlay may help in warning people of hazardous areas or dangers.
Losses	Remote working using XR potentially leads to less physical contact with humans, which may lead to people feeling more lonely and less integrated. For some people it can be more challenging to wear a headset, especially over a longer time duration, resulting in pain in the head, neck, back, etc. The consequences from looking at a short distance for a longer time are still unknown. The XR overlay may impair the human's judgement of hazardous situation.
Key value indicators (factory)	Working accidents and occupational diseases of factory personnel. Number of self-reported mental-health illnesses among factory personnel. Self-reported feeling of isolation and feelings of alienation among factory personnel.
Key value indicators (society)	Reported working accidents and occupational diseases Number of self-reported mental-health illnesses potentially related to XR Self-reported feeling of isolation and feelings of alienation

Income and work

Key impact areas:	Unemployment, employment, prevention of physical safety issues, promotion of safe and secure working environments, facilitation of reskilling and upskilling possibilities, inclusion at workplaces, promotion of decent working hours
Gains	New job opportunities for people who can get trained remotely with the help of XR.
	Remote working via XR may allow workers to avoid having to enter unsafe or insecure working environments.
	XR may facilitate reskilling and upskilling possibilities thereby reducing unemployment rates.
	XR allows to adapt the XR content to the individual worker, thereby opening up for older people/disabled people to be included to a larger extend.
	Remote XR expert support permits more decent work hours for the experts as they do not need to travel.
Losses	People not being able to work with XR are left out of the job market.
	The XR overlay may impair the human's judgement of hazardous situation.
Key value indicators (factory)	Subjective perception and evaluation of personal employment situation of factory personnel. Average exit age from the labour market for factory personnel. Number of reported hazardous incidents in factory.
Key value indicators (society)	Unemployment rate (total, male, female, age groups) Employment rate in knowledge-intensive sectors Average exit age from the labour market

Social equality and inclusion

Key impact areas:	Equal opportunities, inclusion
Gains	Remote support and overlay in XR device may allow for less skilled workers or workers with disabilities to perform tasks on the factory floor.
Losses	When the capability to be able to handle an XR device becomes a pre-requisite for being employed this could lead to a clear disadvantage for people with less favoured background not being able to afford an XR device on their own.
Key value indicators (factory)	Diversity among factory personnel.
Key value indicators (society)	Unemployment rate (total, male, female, age groups, other groups)

Privacy and integrity

Key impact areas:	Balance between transparency, traceability, and privacy Integrity of data
Gains	-
Losses	XR devices can capture visual and auditory information both from the XR user him/herself but also from the environment of the XR user without consent.
	Sensitive information can be inferred or estimated about the XR user or the environment of the XR user without consent.
	Leaked data from the XR device can cause significant harm.
	Manipulation of the XR content shown to users can cause significant damage.
Key value indicators (factory)	Subjective perception and evaluation of personal understanding of privacy and integrity. Data integrity assessment.
Key value indicators (society)	

4.3.2 Environmental sustainability dimension

Key impact areas:	Emission of greenhouse gases, material resources, land use
Gains	XR enables remote working thereby reducing CO2 emissions due to reduced travelling.
Losses	The production of XR devices is connected with using material resources, land use, emission of greenhouse gases, energy, etc.
Key value indicators (factory)	Reduction in travelling when using XR devices. Lifecycle assessment and utility of XR devices in use on the factory floor.
Key value indicators (society)	CO2 emissions, etc.

4.3.3 Evaluation of XR in manufacturing

The KVI analysis of the use case "XR in manufacturing" shows that the use case is expected to have impact in a number of different areas. The main gains are found in the areas of education and work/income which the use case was designed to improve. Several indirect gains are related to the possibilities of remote working/training. The main losses are found in the areas of healthcare, privacy and integrity. Ensuring a value driven design of the use case means working towards mitigating these losses from the start.

DETERMINISTIC6G works towards supporting offloading functionality to the cloud. Thereby the XR device can become lighter and less bulky, thus mitigating the loss for health issues related to wearing the headset. DETERMINISTIC6G is furthermore working towards supporting more complex rendering etc. for a more immersive experience. While this improves the overall perception it may at the same time lead to more difficulties in understanding the difference between what is real content and what is overlayed content. DETERMINISTIC6G's work towards developing E2E security mechanisms, starting with the assessment of the security principles, see Section 5, is one step towards the mitigation of losses in the area of security and integrity.

In a KVI analysis taking in all sustainability dimensions also the impact of producing XR devices needs to be considered since they are not part of all factory floors today. If the assumption is that the XR device is applied in more than one use case, this has however to be reflected on in the overall sustainability assessment.

4.4 KVI analysis of Exoskeleton in Industrial Context

In the following section, the KVI analysis for the use case of "Exoskeleton in Industrial Context" is performed. For a more detailed description of the use case see Section 3.3. Tables are provided for all impact areas where important gains or losses have been found. The impact areas with respect to Table 4.2 and Table 4.3 Table where no gains or losses have been found are left out.

4.4.1 Socio-economic sustainability dimension

Connectivity

Key impact areas:	Access to affordable, high quality, reliable, sustainable, and resilient ICT
Gains	Remote monitoring through digital twin of the system requires connectivity at more remote places, this will be a gain for remote areas to get better connectivity.
Losses	-
Key value	Digital inclusion (reflecting SDG #10. Reduced inequalities).
Key value indicators (factory)	-
Key value indicators (society)	Proportion of households with access to good connectivity.

Healthcare

Key impact areas:	Health, reduction of injuries
Gains	Reduction of adoption barriers.
	Adaptive physical support for demanding tasks.
	Ability to conduct on-line ergonomics assessment and intervention using technologies like digital twin.
Losses	-
Key value	Societal sustainability (cost reduction for the care of work-related injuries) Personal health and protection from harm (reduction of number work-related injuries)
Key value indicators (factory)	Working accidents and occupational diseases of factory personnel.
Key value indicators (society)	Cost reduction for the care of work-related injuries. Reduction of number work-related injuries

Income and work

Key impact areas:	Unemployment, employment, prevention of physical safety issues, promotion of safe and secure working environments, facilitation of reskilling and upskilling possibilities, inclusion at workplaces, promotion of decent working hours
Gains	Inclusion of broader workforce to execute tasks requiring more force.
	Prevent physical safety issues, promote safe and secure working environments (reduce risk of environmental related hazards).
Losses	-
Key value	Digital inclusion (reflecting SDG #10. Reduced inequalities). Simplified life (reflecting partly SDG #3 Good Health and Well-being, #9 Industry, Innovation and Infrastructure). Personal health and protection from harm.
Key value indicators (factory)	Reduction of number work-related injuries.
Key value indicators (society)	Reduction of number work-related injuries.

Social equality and inclusion

Key impact areas:	Ensure equal opportunities
Gains	Cost reduction of the hardware of the device, due to the offload computation can lead to a increased spread of the technology.
Losses	-
Key value	Accessibility to the technology (reflecting SDG #10 Reduce inequality).
Key value indicators (factory)	Number of factories that have adopted the technology.
Key value indicators (society)	Number of users who have adopted the technology.

Privacy and integrity

Key impact areas:	Balance between transparency, traceability, and privacy Integrity of data
Gains	Enhanced data integrity due to more robust real-time data processing capabilities.
Losses	Potential risk of data leakage during the transfer of data from the exoskeleton to the external computing infrastructure.
Key Value	Trust. Privacy and confidentiality.
Key value indicators (factory)	Reported confidence in the device. Absence of data breaches or unauthorized access incidents.
Key value indicators (society)	-

4.4.2 Environmental sustainability dimension

Key impact areas:	Energy, material resources
Gains	Reduced use of battery due to a reduced power consumption (delocalization of components).
	Predictive maintenance thanks to digital monitoring of the system.
Losses	Spreading of exoskeleton may lead to a resource consumption. Exoskeletons require significant amounts of energy, materials, and resources to manufacture, operate, and maintain. This can result in a high environmental impact, especially if the materials used are non-renewable or difficult to recycle.
Key value	Environmental sustainability.
Key value indicators (factory)	-
Key value indicators (society)	Number of exhausted batteries.

4.4.3 Evaluation of Exoskeleton in industrial context

The KVI analysis of the use case “Exoskeleton in Industrial context” examines various factors related to the implementation and utilization of exoskeleton technology in industrial settings. It focuses on assessing the Key Value Indicators (KVIs) that are relevant to this specific use case. The analysis involves evaluating the impact of exoskeleton adoption on factors such as productivity, worker safety, ergonomics, and overall operational efficiency. It considers how the integration of offload computation capabilities in the exoskeleton technology can further enhance its performance and effectiveness. Through the KVI analysis, the potential benefits and challenges associated with implementing exoskeletons in industrial environments are identified and assessed. This includes understanding the potential cost savings, improvements in worker well-being, and the potential for enhanced task performance and accuracy. It takes also into account potential drawbacks of the solution, referring to possible data leakages, degradation of performances due to the network. It also considers the potential environmental impact of the spreading of the technology that can lead to resource consumption. Moreover, the analysis also examines the compatibility of the exoskeleton technology with existing industrial processes and systems. It explores how the integration of exoskeletons with other technologies and digital platforms can enable real-time data monitoring, analysis, and optimization.

Although not explicitly mentioned, the analysis takes into account the potential reduction of weight and size of the exoskeleton due to the offload computation enabled by DETERMINISTIC6G. The potential weight reduction can have several benefits, such as improving user comfort and mobility, minimizing physical strain, and increasing overall user acceptance and adoption. A lighter exoskeleton can also contribute to enhanced energy efficiency, allowing for longer operation times and reducing the need for frequent recharging or battery replacements.

4.5 KVI analysis Adaptive Manufacturing

This section shows the KVI analysis for the to the Adaptive Manufacturing use case. More information about the use case itself and the detailed use case scenarios can be found in Section 3.4.

4.5.1 Socio-economic sustainability dimension

Energy

Key impact areas:	**Upgrade technology for supplying modern, sustainable energy systems**
Gains	Improved availability of energy technologies (batteries, photovoltaic panels, technology for wind turbines)
Losses	Increased energy usage due to higher number of devices and increased required computation power. This can negatively impact the availability of energy for the surrounding environment. Energy rationing can furthermore impact the private population.
Key value	Affordable energy systems, energy usage
Key value indicators (factory)	Number of affordable energy production and storage systems
Key value indicators (society)	Number of affordable energy production and storage systems

Water & Sanitation

Key impact areas:	**Increased water-use efficiency**
Gains	Mobile processing modules can be cleaned with better optimization on their previous usage and the resulted contamination. The usage of water and other cleaning supplies can be adapted to that.
Losses	-
Key value	Water use efficiency
Key value indicators (factory)	Amount of used water for cleaning processes
Key value indicators (society)	-

Food

Key impact areas:	**Improve agriculture productivity, improved knowledge**
Gains	Improved adaptation of processing steps based on properties of incoming source material
	Mobile processing modules allow interactions like cleaning, maintenance, repairs to be done away from the production line. • Less pollution of the rest of the machine • Clean room
	Reduced costs of food production via • Reusing processing modules for multiple production lines, leads to reduced hardware cost • Dark factory: increased amount of fully automated processing steps, fewer or no human personnel required
Losses	-
Key value	Agriculture productivity
Key value indicators (factory)	Efficiency of agricultural manufacturing or processing
Key value indicators (society)	Reduced cost of food

Healthcare

Key impact areas:	Access to quality healthcare services
Gains	Reduced cost via more efficient production
	Reduced cost via easier establishment of clean rooms / clean areas
	Reduced wasted resources via adapted processing
Losses	-
Key value	Cost of medicine
Key value indicators (factory)	Reduced production cost of medicine
Key value indicators (society)	Reduced cost of medicine

Income & Work

Key impact areas:	Prevent physical safety issues, Employment
Gains	Improved safety via determinism of used communication. Essential for collaborative machines with mobile devices (AGVs, MPMs)
Losses	Dark factory: increased amount of fully automated processing steps, fewer or no human personnel required
Key value	Worker safety
Key value indicators (factory)	Number of workplace accidents with involved mobile devices
Key value indicators (society)	-

Well-being & Culture

Key impact areas:	Improved rights to practice one's culture
Gains	Individualization of products simplifies the provision of tailored products for smaller social or cultural groups
Losses	-
Key value	Cultural representation
Key value indicators (factory)	-
Key value indicators (society)	Availability of individualized products for cultural representation

Social equality & Inclusion

Key impact areas:	Empower and promote the social, economic and political inclusion of all
Gains	-
Losses	Machines will mainly be owned by rich societies, which intensifies the economic inequality even more.
Key value	Social and economic inequalities
Key value indicators (factory)	-
Key value indicators (society)	-

4.5.2 Environmental sustainability dimension

Emissions to water, air and soil

Key impact areas:	Reduction in emissions
Gains	Localized production of goods
Losses	-
Key value	Emissions caused by transport
Key value indicators (factory)	Goods and products can be produced closed to the customer?
Key value indicators (society)	-

Land use

Key impact areas:	Reduce land footprint
Gains	Reduced factory size due to increased flexibility and modularity of production lines
Losses	-
Key value	Land usage
Key value indicators (factory)	Reduced land usage for factories
Key value indicators (society)	-

Material resources

Key impact areas:	Improve resource efficiency in consumption or production
Gains	Reduced need for materials by reusing the same processing modules for multiple production lines.
Losses	-
Key value	Material usage for machines and factories
Key value indicators (factory)	Reduced number of required devices
Key value indicators (society)	-

4.5.3 Evaluation of Adaptive Manufacturing

The KVI analysis of the Factory Automation: Adaptive Manufacturing use case resulted in different affected areas that are mostly caused by the key aspects of Adaptive Manufacturing. These are the increased flexibility and modularity of factories. These increments improve the efficiency of different factory processes and the efficiency of factories themselves. These efficiency improvements can further lower the final costs and therefore increase the affordability and accessibility of different goods, products, etc. This includes food, medicine or devices for energy production and storage.

The identified potential losses generally arise from the nature of automation, in particular factory automation. This is for one the omnipresent potential for loss of factory related jobs, but also the implicit inequality regarding the affordability of new and improved machines, or factories in general.

The identified gains as well as the identified losses can be used as reference for further development of different applications within the field of Adaptive Manufacturing using new technologies like 6G to focus on generating improvements for the society.

4.6 KVI analysis Smart Farming

The following sections highlight the sustainability, environmental and socio-economic impact of smart farming.

4.6.1 Environmental sustainability

Key impact areas:	Reliance on natural energy resources and reduction of wastage
Gains	By providing deterministic communication, 6G can support applications regarding power management in farms, enabling them to use natural energy resources more efficiently.
	Efficient utilization of water and soil is possible through use of automation in the farms
Losses	-
Key value indicators (factory)	Energy bills
Key value indicators (society)	-

4.6.2 Socio-economic sustainability dimension

Food and Resilience

Key impact areas:	Food wastage
Gains	Identifying and preventing loss of crops due to infestation, adverse weather, pest etc., by using continuous monitoring and AI
	Fast agricultural operations such as harvesting and post processing of the yield by use of automation
Losses	-
Key value indicators (factory)	Amount of yield before and after automation
Key value indicators (society)	Average food prices

Income and work

Key impact areas:	Prevention of physical safety issues, promotion of safe and secure working environments, inclusion at workplaces, promotion of decent working hours
Gains	Using automated smart farming processes, farmers don't have to spend long hours working in the field or perform hazardous operations
	Using automation, the farmers can work normal working hours
	Using automation, inclusiveness can be fostered by involving disabled and elderly people
Losses	Less manual labour needed, meaning a loss of jobs for certain groups of people
Key value indicators (factory)	Evaluation of personal employment situation Average exit age from the labour market (factory)
Key value indicators (society)	Unemployment rate (total, male, female, age groups) Employment rate in knowledge-intensive sectors Average exit age from the labour market

4.6.3 Evaluation of Smart Farming

Using automation and the deterministic qualities of 6G, it will be possible to support autonomous applications which will improve the quality of life of farm personnel, agricultural yield and the efficient use of natural resources. With a high degree of automation, required manual labour will be reduced, but it will be possible to have more inclusive employment.

4.7 Conclusions on KVIs

The KVI analysis of the use cases presented in this section shows that it is possible to gather indications for evaluating the potential benefits and drawbacks of use cases upfront via key values and key value indicators. There is however a need to reduce the complexity to allow for feasible first evaluations. Furthermore, many of the KVIs can only fully be evaluated once the use case is implemented. Therefore, a continuous evaluation both by subject matter experts as well as individual assessments is needed throughout the entire process of designing, implementing, and updating the use cases in manufacturing. Identifying enablers to maximize gains and minimize losses and understanding drivers and barriers for different potential impacts is a crucial step.

5 Principles of security assessment

5.1 Introduction

To provide the key ingredients described in Section 1.1, i.e., adaptability, confidence and reliability, one of the key aspects is guaranteeing the security of the communications. The security level of the network communications needs to be adapted to the requirements of the applications. It is always necessary to consider a trade-off between the security mechanisms, the risks and the performance. Furthermore, end-to-end determinism needs to be protected from both malicious and unintentional denials of service that are in many cases easy to perform. This requires the deployment of security monitoring and controls that will guarantee the confidence and reliability needed by the applications.

This section describes the results obtained from gathering and analysing use case requirements for a new security-by-design deterministic networking deployment. Their goal is to extract specific security needs, constraints, vulnerabilities, and countermeasures tailored to the unique requirements of each use case to obtain a security architecture for DETERMINISTIC6G. The focus is in ensuring a robust security framework without reiterating the general security issues already covered by DetNet and TSN 3GPP/ETSI/IEC-IEEE standards.

The diverse use cases in the project (i.e., XR, exoskeletons, factory, and mobile automation) that use or plan to use deterministic networking, participated in discussions and filled out a questionnaire-table that helped identify specific security concerns associated with each use case. In this way, it was possible to determine how security functions should be designed so that they offer prioritised, adapted, and optimised solutions to protect the confidentiality, integrity, and availability of deterministic networking. The questionnaire-tables included the following topics: system characteristics, trade off strategies, potential use of security SLAs, identified vulnerabilities, required security functions, prioritisation of the risks from different types of attacks, and possible types of responses. End users are not familiar with all these issues, but the insights gathered have led to a better understanding of common issues and differences between use cases and even between different scenarios within each use case.

From this study it is possible to elaborate the tailored security functions and extensions that form part of the deterministic network architecture, meet the unique needs of different industries, and obtain a robust and reliable 6G communication infrastructure that will not compromise critical applications and data.

5.2 Brief description of security in existing standards

As indicated above, many of the security mechanisms are already covered in standards but integrating novel architectural concepts to achieve improved predictability and determinism for wireless communication introduces new vulnerabilities and requirements. It is also necessary to consider the integration of an end-to-end security system, security-by-design concepts, and "delay transparency". The latter refers to the need to consider the additional delays due to security measures that should be included in the accepted delays.

First, we have the TSN profiles for Industrial Automation that include standards developed by IEC, IEEE and IETF. These standards are concerned with the security between stations, shared security, security of the security functions, and reuse of security mechanisms. These concern the industrial context but leave out the application and middleware components. As can be seen the Figure 5.1.

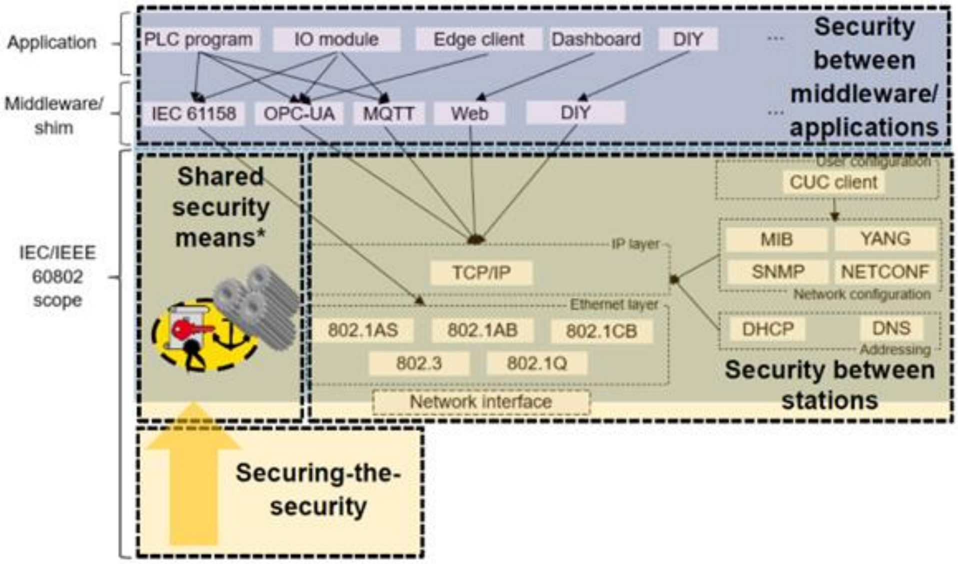


Figure 5.1: IEC/IEEE 60802 Security (TSN profiles for Industrial Automation) [FIS21]

For the security between stations, the protection of messages, authorisations and the TSN management in general are included. For the shared security, encryption, key management, and the definition of device identifiers are included. For the reuse and adaptation of existing mechanisms, we have the requirements defined by 802.1AR, NETCONF for the configuration and supervision of TSN nodes, trust anchors and credentials for Industrial Automation components, and other 802.1 security standards.

In DETERMINISTIC6G, we will not deal with all these aspects that are already defined or ongoing such as encryption techniques to use, configuration, credentials, etc. except to see how they could impact determinism and low latency or be used for managing and assuring truly deterministic and low latency network flows and determine if any recommendations or best practices can be elaborated for this specific goal.

5.3 Security requirements concerning deterministic networking

In deterministic networking, network behaviour is expected to be predictable and consistent, to provide reliable and deterministic communications for time-sensitive applications. Concerning security, besides the vulnerabilities that exist in standard IT/OT systems, assuring the latency and determinism present extra challenges even more acute when wireless communications are used (e.g., IoT, 5G/6G networks). The main security concerns include:

- Attacks affecting the Time Synchronization [Bus23]: Disruption of time synchronization protocols will desynchronise network devices and may lead to performance degradation or service/system failures.
- DoS attacks [Sha23b]: Flooding the network with network traffic or requests, or exploiting vulnerabilities in the deterministic protocols, will cause service degradation or disruptions. Even DoS attacks which make use of low-bandwidth rate to accomplish their purpose (i.e., Slow DoS attacks) can negatively impact determinism.
- Unauthorized Access: Even though this is also an issue in IT networks, deterministic networking may require the enforcement of strict policies to prevent unauthorized access to

critical deterministic network components without negative impact on the expected properties.

- Attacks on multi-access edge computing (MEC) [Gya22]: Distributed computing architectures that bring computational resources and services closer to the network edge can help achieve low-latency and high-bandwidth and reduce the amount of communicated data. The use of MEC can introduce vulnerabilities but can also improve the security by allowing better control of the confidentiality, integrity, and availability (CIA) of the data, e.g., localized threat detection, encryption, and access control.
- Attacks on data integrity and confidentiality: Prioritising performance and determinism over data integrity and confidentiality, as often done by deterministic or low-latency applications, may pose security risks. Sensitive data needs to be protected to avoid eavesdropping, tampering, or gaining unauthorized access to this data.

Other challenges need to be considered even though they are generic to all types of networking. The negative impact on determinism needs to be considered, as well as how to prevent or mitigate them. These challenges are, for instance, the vulnerabilities introduced by end-to-end convergence of the different technological domains involved, the continuous adaptation between the application and network layers, and the use of ML techniques for detection, prediction and decision control.

To address these security concerns, the DETERMINISTIC6G project will define a *security-by-design* architecture that allows implementing robust security measures such as strong authentication mechanisms, secure communication protocols, intrusion detection, continuous high-precision monitoring, and remediation strategies. The goal is to ensure network reliability, integrity, and confidentiality while maintaining the desired deterministic behaviour required by different use cases.

For this, it is necessary to incorporate security enablers at all levels of design and implementation of 5G/6G networks, including: secure communication protocols (e.g., Transport Layer Security); authentication and access control to critical components and assets of deterministic applications; threat detection and intrusion prevention that monitor network traffic and identify deviations from expected behaviour, and trigger remediation actions; secure time synchronization by ensuring the integrity and authenticity of time protocols; security of MEC; secure network slicing and software-defined networking (SDN) for separating different types of devices and applications based on their security requirements; providing redundancy to ensure high availability and resilience; monitoring for high-precision measures and feature extraction; monitoring and logging for forensics analysis; systematic security testing and validation to identify any vulnerabilities.

Other aspects that are not covered by DETERMINISTIC6G are securing firmware and software updates, security training of users, operators and providers of vertical solutions.

The first step for specifying the security-by-design architecture was to analyse each use case to identify the requirements coming from each. Then a synthesis allows determining the common security requirements, the specificity introduced by determinism, and the differences to identify the level of adaptability and dynamicity of the security mechanisms.

To analyse the requirements as they are perceived by the use cases, discussions were held, and a table was filled out for each use case describing the issues involved for each security issue and system characteristics. From this, besides the security concerns and enablers listed above, the following aspects have emerged and need to be considered. The use cases analysed were XR, exoskeletons, factory and mobile automation.

The main aspects to consider are:

- Security functions must be flexible in order to adapt to the different requirements coming from each use case but also internal to each use case. Different use cases and data in each use case require different levels of security.
- Safety is crucial in many use cases meaning that the systems are cyber-physical systems where the impact of one on the other needs to be considered.
- The tendency is to shift from completely closed systems to open ones that can be accessed from the exterior (e.g., through web applications), involve multi-providers, use private or public cloud services, involve non-managed devices and applications, integrate open-source software, use of wireless communications (e.g., IoT networks, 5G/6G networks).
- Time synchronisation is crucial, but the level of tolerated deviations depends on the use case. In general, the synchronisation protocol exchanges do not need to be encrypted but the integrity and the authenticity of the packets needs to be assured.
- The confidentiality, integrity, and availability (CIA) needs to be assured but depends on the needs of each use case scenario.
- Security requirement can change depending on the state of the applications and environment (e.g., emergency situations, degraded operation mode).
- In some cases, multi-providers and multi-tenants need to interoperate and this should include the negotiation of security properties.
- Root cause analysis and forensics could be necessary in some use cases.
- Encryption of data and control plane exchanges is needed in most applications to prevent unauthorized access. Sensitive data needs to be identified, encrypted and stored securely to prevent unauthorized access and disclosure.
- Redundancy: In case of a network failure or other disruption, the exoskeleton should have redundant control mechanisms in place to ensure that the user can still operate it safely and the service continuity is assured.
- DoS and Slow DoS attacks need to be blocked before they impact the services.
- Depending on the use case, packet loss, jitter, synchronization accuracy, and communications may need to be monitored to detect any deviations from expected behaviour.
- The complete automation of the security management may be necessary, but, in some cases, human interventions might be necessary or preferred to identify the nature of a security breach and determine how to deal with it.
- The use of wireless communications can be vulnerable to jamming attacks, fake devices and base stations.
- In applications that depend on continued flow of data, such as vehicle positioning and obstacle detection, encryption is not always possible and can allow attackers to disrupt them by injecting/replaying packets and requests.

5.4 Summary

From the previous analysis, we have the following main aspects that need to be considered in the design, implementation, and operation of security for deterministic networks adapted to the different use case requirements:

- There is a need to distinguish the security required by the different scenarios and assets. For this security policies need to be defined in a way that can be easily understood by the operators and users of the vertical applications. This can be done by formally defining Security Service Level Agreements (SSLAs) that indicate what security controls and remediation strategies need to be deployed and how to configure the security functions. In addition, it should be possible to define different security levels that would allow to dynamically change the security position depending on the state of the system and its environment. These SSLAs can be static (defined at the beginning and not changed) or dynamic (can be changed during operation). They are also needed to enable the interoperation with other systems (e.g., environments that involve multi-providers, multi-tenants and/or multi-domains: radio, core, edge…).
- Special attention should be paid to prevent disruption of wireless communications (e.g., caused by jamming attacks, fake base stations and devices, artificial noise injection) since this can impact many of the applications using IoT and wireless communications.
- The type of attacks that can have the most negative impact on determinism are DoS and Slow DoS. These are particularly easy to exploit by attacks and are difficult to detect when done intelligently, and difficult to prevent before they affect the targeted services.
- Continuous monitoring is required for high-precision measurements, real-time detection of breaches and response, feature extraction for AI-based analytics (e.g., for behaviour analysis, optimisations, decision support and root cause analysis), and information logging for forensics. To detect anomalies, multimodal data monitoring (e.g., of protocols, business activity, behaviour, location, KPIs such as bandwidth, latency, jitter, packet loss…).
- Network segmentation based on secured slicing should be considered to isolate certain interactions, separate DetNet/TSN streams and non-DetNet/TSN traffic, dedicated slicing for safety re-enforcement, implementing layers of defence, replication, and moving target defence strategies.
- Protecting AI/ML techniques is necessary when they are used by making them explainable, and detecting manipulations of the training data and models, etc.
- Protecting the use of virtualisation, e.g., Software Defined Networking (SDN), Network Function Virtualisation (NFV), MEC, private and public clouds. For Cyber-Physical systems or systems that can affect the safety of humans and assets, it is necessary to consider the impact of the physical domain to the cyber domains and vice versa.
- For optimisation of the security mechanisms, there is a need to define trade-offs between the estimated risk, the cost of the security mechanisms and the resulting performance, e.g., between the security level and resulting KPIs (e.g., latency, jitter, determinism), between the required encryption and resulting latencies, etc.
- Different functions or aspects required by the deterministic networks need to be assured, such as, time synchronisation, traffic aggregation (non-TSN and TSN traffic),
- End-to-end security analysis and prevention/remediation is needed, as well as zero-touch network and service management (ZSM) and orchestration for E2E response to security breaches. The management of security breaches (e.g., response, mitigation actions,

reconfiguring the network) should be automated but also allow human interventions in some cases.

- Another consideration that could be important in some applications is forensics related to deterministic and latency for investigating security incidents, identifying root causes of attacks, and gathering information for legal or regulatory purposes.

6 Computing resource requirements

6.1 Introduction

This section provides a set of architectural and functional requirements for the compute (e.g., Edge cloud) deployment to ensure real-time operation. Furthermore, a detailed description of computing resource requirements for the use cases from process offloading perspective is also provided.

6.2 Generic deterministic cloud requirements

Deterministic applications with specific time-aware requirements concerning predictable or bounded latency, low jitter as well as high reliability, etc., require special features provided by the compute infrastructure, the virtualization solution, and the runtime environment.

The requirements discussed in this section are derived from the use cases presented above in this document and mainly covers generic system requirements, basic infrastructure requirements, considerations on 3GPP System to Edge infrastructure interfaces and APIs as well as reliability and performance requirements.

Basically, the cloud infrastructure should provide:

- Scalable, automated compute services including the support of various runtime execution environments, such as Virtual Machine (VM), containers as well as serverless runtime.
- High performance and a scalable environment for supporting the deployment of Virtualized Network Functions (VNFs), including an automated orchestration and management layer, covering the whole workload lifecycle management.
- Capabilities to host real-time operating systems to provide timing guarantees for the execution of application processes.
- For containerized applications, adequate isolation between the containers and the cloud resource management needs to ensure the required CPU resources for the containerized application instances. The container network interface plugins should have the capabilities to ensure time-aware traffic handling on container network level. The container management system must be capable to ensure dedicate resource usage for time-critical application processes, such as dedicated hardware, CPU pinning, etc.
- Infrastructure or application accelerators (e.g., GPUs, Smart Network Interface Cards (NIC), persistent memory) in the cases of time-critical and compute intensive applications, to guarantee the required quality for the communication and the application services and may also be used to offload workload from the CPU. If these special hardware capabilities are available, these ones should be advertised towards the virtualization domain to the application and its management system.
- Support for IEEE Time Sensitive Networking (TSN) and/or IETF Deterministic Networking (DetNet) services. This would require specialized, TSN capable NICs, to guarantee bounded latency, time synchronization support on the hardware and software (e.g., capability to synchronize containerized applications) layers. Native handling of layer 2 (Ethernet) traffic handling should also be ensured in the virtualization domain considering VM- or container-based application deployment options.
- Programmable and scalable storage services.

- The infrastructure – including the hardware and software layers – may be built up from components belonging to different vendors, hence multi-vendor interoperability must be ensured.
- Different service models, such as Infrastructure as a Service, Platform as a Service, Software as a Service should be supported depending on the application requirements.
- Multiple mechanisms to separate different domains, such as virtualized communication, end-user applications, runtime execution, management, etc.

To improve the reliability of services, the cloud infrastructure should additionally provide:

- The capability of hosting multiple application instances using redundant resources. This may require server (data center) hardware redundancy, including redundant connectivity and networking and special levels for power supply.
- Support for the required reliability and isolation features by the cloud management system during the application deployment phase, e.g., by enforcing the affinity/anti-affinity rules in the placement process, specific resource application schema as well as during the whole application runtime.
- The cloud management should also ensure seamless, automatic fault handling strategies, e.g., restoring the failed virtualized communication infrastructure and application components. Optionally, the restoration of the application components could be ensured by the application management; in this case the cloud infrastructure must provide the necessary configuration capabilities.

Basic interface API considerations:

- The interfaces to expose the 3GPP Core network functions for control QoS requirements and influence the traffic steering as specified in TS 23.501 [3GPP23-23501], TS 23.502 [3GPP23-23502] as well as in TS 23.548 [3GPP23-23548] should be supported.
- The interfaces to enable the handling of Edge Application Servers (e.g., registration, discovery, provisioning) and to provide exposure APIs towards the application clients as specified in TS 23.558 [3GPP23-23558] should be supported.
- ETSI MEC-based API for infrastructure management (e.g., deployment options for applications, traffic routing, network exposure) should be supported.
- Exposure of management services for 3GPP-enabled Edge computing is required to support according to TS 28.538 [3GPP28-23538].
- Interface APIs towards the TSN and DetNet control plane to explore cloud infrastructure capabilities and services and to configure application deployment details should be provided by the cloud management.

6.3 Requirements on computing offload

Four classes of computing resources can be identified, each of them being relevant in offloading:

- *On-board computing:* computing on end devices.
- *Local and customer edge computing:* computing on a local server, or edge computing resources that are deployed locally by the customer, within its network and IT resources.

- *Network edge computing and Multi-access Edge Computing (MEC):* computing at the edge of a network operator, according to the cloud computing model, for both its network services and customer services.
- *Cloud computing:* cloud service offered on large data centers, according to the traditional cloud service model.

The focus is on use cases directly associated to industrial services (industrial exoskeleton, industrial automation, extended reality), but the approach can be extended to other uses cases as well (mobile automation: smart farming, ...). For each of the use cases, computing requirements can be identified, described, and characterised in each of the computing resource classes whenever it is applicable. This information is first used to capture and provide a synthetic view of computing resource that are involved in each of the use cases. Then each of the requirements can be detailed, for example in terms of timeliness, reliability, performance, etc., possibly with indications on the types of hardware or software mechanisms that can be used. Table 6.1 indicates the typical computing resources to be used for some of the operations involved in use cases.

Table 6.1: Use case computations and adequate computing resources

Use case	On-board	Local and customer edge	Network edge and MEC	Cloud
Exoskeleton	Low-level control	Low-level and Middle-level control	Low-level and Middle-level control	Middle-level and High-level control
Factory Automation	Computing on AGV, MPM	Industrial applications AGV, MPM computation offload	Industrial applications AGV, MPM computation offload	Scenarios combining Robotic AGVs, XR and Exoskeletons
XR	Device computations	Rendering Spatial compute Application logic	Rendering Spatial compute Application logic	Application logic
Mobile Automation	Computing on UAV, UGV, robot, sensors, meters	UAV, UGV, robot computation offload; gateways; image processing.	UAV, UGV, robot computation offload; image processing, microgrid control.	Control center, monitoring and remote control, image processing, microgrid control.

6.4 Common elements

In many cases computing performance is essential, a key element influencing the capacities of the infrastructure, and the applications. As basic performance enhancement tools, dedicated hardware technologies and accelerator technologies such as GPUs, FPGAs, or multicore processors, can be used in almost all computing resources, and for almost all applications. However, these dedicated hardware and accelerators are generally used to complement general-purpose processors (GPPs), which is a source of additional system costs, and additional complexity in the infrastructure. An operating system (OS) is required to be installed and running on each GPP. In most cases, real-time and open OS may be preferable (real-time Linux-like) to facilitate integration and interoperability, and the realization of deterministic communication services. Real-time constructs in an OS are usually a key element for reliability and latency guarantees, which are almost always important pillars in deterministic communications. For servers running on edge computing, MEC or cloud environments, the standard configurations and operations of such environments requires a virtualization software infrastructure (VM-based or container-based for cloud-native environments), in addition to the real-time OS. Of course, the use of special purpose hardware and accelerator is also possible, but it should be kept in mind that edge computing and cloud computing environments are normally to be shared between several organizations, applications, and users, which can make deterministic services more difficult to realize.

6.5 Exoskeleton computations

- *On-board exoskeleton computing:* The on-board system of an exoskeleton can be implemented on dedicated hardware, possibly with accelerators, depending on the complexity of tasks involved in its main components: sensors, actuators, motor drivers. In the case of a near-term scenario, the low-level control system is on board, and its implementation can be envisioned with embedded GPP, running an embedded real-time OS.
- *Local and customer edge exoskeleton computing:* The smart factory in an industrial exoskeleton system, which is typically used for environmental and task-related sensing system, should be implementable with local and end-user edge computing resources, typically organized and operated as a smart factory cloud, an edge computing system for the factory. Resources in such a system can be composed of dedicated hardware, accelerators, and servers based on GPP. The software composition is that of an edge cloud environment, with hypervisor(s) and/or container platform(s), on which real-time OS(s) is(are) installed.
- *Network edge computing and MEC exoskeleton computing:* The smart factory in an industrial exoskeleton system should be also implementable with edge computing resources at the edge of operator networks, which can be typically the operators of MEC resources (but not only). Network edge computing and MEC resources can be allocated to host a smart factory cloud, especially in case the industrial exoskeleton company does not have internal resources to build its own factory cloud. Computing resources in such a system can be composed of dedicated hardware, accelerators, and servers based on GPP. The software composition is a shared cloud environment, with hypervisor(s) and/or container platform(s), on which real-time OS(s) is(are) installed. The appropriate sharing and security of such a platform is highly important, which needs the implementation of adequate cloud management and security software.
- *Cloud computing for exoskeleton:* In the near-term scenario, parts of the exoskeleton control system are offloaded to the cloud. In the long-term scenario, the totality of the control system is on the cloud. This supposes that the cloud environment in use is dimensioned accordingly, and it is adaptable to support complex exoskeleton control strategies. Dedicated hardware, accelerator, servers based on GPP, and storage systems are generally required for such a system. In terms of software, the compositions of edge computing and MEC environments are also valid here. Also, to enable to use such a cloud environment, real-time cloud services and deterministic end-to-end communications between the cloud and on-board skeleton computing are required.

6.6 Factory automation and adaptive manufacturing computations

- *On-board industrial computing:* On-board computing is found in AGVs, safety systems (cameras, …) or MPMs components of the adaptive manufacturing scenario. Moreover, on-board computing is present almost everywhere in the adaptive manufacturing scenario combining collaborative robots, extended reality, and exoskeletons. These on-board computing services can also be implemented on dedicated hardware and accelerators. Some of these on-board computing should be implementable with embedded GPP and real-time OS, an alternative that could appear particularly interesting for adaptive manufacturing.
- *Local and customer edge industrial computing:* These computing resources are central in running industrial automation applications, and the use of a factory cloud build on costumer edge computing resources could be essential for adaptive manufacturing computations, in which the amount of computing resources is dynamically adapted to the needs of adaptive manufacturing computations. Resources in such an edge cloud can be composed of dedicated hardware, accelerators, and servers based on GPP. The software composition is that of an edge cloud environment, with hypervisor(s) and/or container platforms, on which real-time OS(s) is(are) installed.
- *Network edge computing and MEC for industrial applications:* Adaptive manufacturing computations with AGVs and MPM components, and the combined scenario with collaborative robots – extended reality – exoskeletons, can be also implementable with edge computing resources at the edge of operator networks and MEC resource. Network edge computing and MEC resources can be allocated to an adaptive manufacturing system, especially in case the manufacturing company does not have internal resources to build its own factory cloud. Computing resources in such a system can be composed of dedicated hardware, accelerators, and servers based on GPP. The software composition is a shared cloud environment, with hypervisor(s) and/or container platforms, on which real-time OS(s) are installed. The appropriate sharing, security, and dynamic resource allocation of such a platform are key to success, which needs the implementation of adequate cloud management, security and dynamic resource scaling software.
- *Cloud computing for industrial applications:* The most relevant scenario involving cloud computing resources, in addition of other (edge computing, on-board computing) resources, is probably the combination of collaborative robots with extended reality and exoskeletons. In this scenario, a hierarchical system of computing resources may help to accommodate the various computing needs in terms of performance, real-time (latency), reliability, end-to-end deterministic communication requirements, flexible computing resource scaling as a foreseen requirement for adaptive manufacturing. Dedicated hardware, accelerator, servers based on GPP, and storage systems are required. The typical software composition of edge computing and MEC environments is the same as in previous sections. Real-time cloud services, elastic and adaptable cloud environment, and deterministic end-to-end communications between the cloud and on-board computing systems are key requirements.

6.7 XR computations

- *On-board computing on XR devices:* Basic device computations (sensor signal, display, rendering…) in general can be done on-board. In many cases, special purpose hardware, dedicated hardware and accelerators may be useful for real-time performance. Spatial compute can be done on-bord, with both dedicated hardware and accelerators, or GPP, on which OS(s) is(are) installed. Also, for OS running on GPP, a real-time and open OS may be preferable (real-time Linux-like) to facilitate the integrations of the device with deterministic network services.
- *Local and customer edge XR computing:* This involves the use of dedicated (local) application server(s) or edge computing server(s) in customer premises, typically for rendering, spatial compute, and the application logic. Special purpose/dedicated hardware or accelerators may also be used in conjunction with GPPs on local application server(s), with an open and real-time OS installed and running on the GPP. In case of servers running on an edge computing environment, a virtualization software is required, in addition to the real-time OS. Edge cloud environments are normally to be shared between several XR applications and users, which can have the potential to make deterministic behaviour more complicated to realize.
- *Network edge computing and MEC for XR:* This involves the use of computing resources at the edge of operator networks, which can be typically the operators MEC resources, and possibly an edge cloud provided resources deployed by a third-party provider to host edge computing applications. The use of these types of computing resources is also well indicated for running the XR application logic. Cloud servers optimized for XR computations may be more appropriate, which may include accelerators, OS and virtualization/container software. Network edge cloud and MEC environments are also intended to be shared between several network services, several XR applications, and several other applications. Thus, ensuring a deterministic behaviour in such an environment may be more complicated to realize. However, the use of this type of computing resources may be appropriate and well indicated for XR application providers that do not have enough computing resources.
- *Cloud computing for XR applications:* For the most complex XR applications, the totality or parts of these applications can be implemented with data center cloud resources. In an XR cloud content streaming scenario for example, data can be distributed between a data center cloud, and several edge computing nodes. In such a scenario, content of an XR application can be processed and streamed from cloud resources, via one or many edge resources, and finally down to the XR devices where on-board computing is made. All types of hardware and software environment sketched in previous sections can be found in a cloud computing environment for XR applications. Deterministic behaviour in a data center cloud environment is even more complicated to realize, but it should be possible to benefit from reliability techniques developed for the operations of such environments.

6.8 Mobile automation: smart farming computations

- *On-board computing on smart farming devices:* Computing capabilities are required on-board the variety of UGV (planters, harvesters, rollers, …) and UAV (drones, aircrafts, …), robot, smart meters and sensor that can be involved in state-of-art and future smart farming asset. Important computing resources may be required, especially for the autonomous operation of UGV and UAV entities incorporating various sensors, sensor data processing, and automated control operations. In many cases, special purpose hardware and accelerators may be used to enable real-time performance. GPP on which real-time OS(s) is (are) installed can also be used to support parts of the autonomous control operations and deterministic communications supports within UGV end UAV, and the other smart farming entities, for the coordination between all these entities during operations.

- *Local and customer edge smart farm computing:* This type of computing resources consists of local application servers, and edge computing resources installed within the premises of the smart farm, typically to provided enough computing resources to run part of the (offloaded) UGV, UAV, robots, and other smart farm entity operations. These resources are also usable to run compute-intensive image processing and related AI analytics algorithms required during smart farming operations. Special purpose and dedicated hardware (for example, for image processing), or accelerator, may also be used in conjunction with GPPs as local application servers and edge cloud servers, with real-time OS installed and running on the GPPs. For the customer edge cloud, a virtualization or container platform software is required, in addition to the real-time OS. In addition to running smart farm entity operations and analytics applications, the customer edge cloud can be used to support the execution of parts of services composing the deterministic 6G network used by the smart farm, for communications and coordination between the entities.
- *Network edge computing and MEC for smart farms:* The use of network *edge computing and MEC* resources is very well indicated for offloading UAV, UGV, and robot computation, especially control and coordination operations that are computationally complex, or parts of remote control of the smart farm entities. These resources can also be used for offloading image processing and related AI analytics, and microgrid control operations for power supply and interconnexion with shared electrical networks. Edge cloud and MEC servers optimized for smart farming computations may be more appropriate, which may include accelerators, OS and virtualization/container software. Network edge cloud and MEC environments are also intended to be shared between several network services, several smart farms and possibly parts of smart grid applications. Thus, ensuring a deterministic behaviour in such an environment may be more complicated to realize. However, the use of this type of computing resources may be appropriate and well indicated for smart farms that do not have enough internal computing resources.
- *Cloud computing for smart farming applications:* For the most complex smart farm applications (monitoring and remote control of smart farm entities and their operations, control center, microgrid control and energy distribution, etc.) the totality or parts of these applications can be implemented with data center cloud resources. Part of each of the combine harvester farming operations, and probably parts of the two other scenario, should be implementable on cloud data centers, a choice that can help in finding enough resources to deal with the complexity of operations. All types of hardware and software environments sketched in previous sections can be found in a cloud computing environment for smart farming applications. Deterministic behaviour in a data center cloud environment is more complicated to realize, but it should be possible to adapt software reliability and fault-tolerance techniques developed for the operations of data center cloud, bringing deterministic communication and computing techniques to cloud-based smart farming applications.

7 Towards deterministic communications in 6G

Considering the emerging requirements from the presented use cases, DETERMINISTIC6G aims to develop a new architecture which will enable 6G deterministic communication. Currently, 5G system architectures from Release 16 onwards support interworking with deterministic communication technology such as TSN and DetNet by introducing several architecture additions, like the novel TSN translator functions at the 5G system boundaries. The 5G architecture and its support for TSN and DetNet are the foundation for the DETERMINISTIC6G system architecture.

In this section, we elaborate on guiding principles for developing the new DETERMINISTIC6G architecture. First, we review the traditional connotation of determinism and then define it in the context of communications.

7.1 A review of determinism

Philosophically, determinism has been defined in several ways. The most common idea is that of *causal determinism* which implies all events in future can be determined completely by past events and the laws of nature [Hoe03]. This idea can be extrapolated to abstract systems, where deterministic systems are those whose behaviour at any point in time is only determined by the input to the system and the rules that govern the evolution of the system. For instance, a deterministic finite-state machine (FSM) is a deterministic system because it is possible to determine the output and the future state of an FSM based on the current state and the input. Subsequently, for a given initial state and a sequence of inputs, the sequence of future states is deterministic, i.e., always bound to happen in the same sequence.

Conceptually, determinism is easy to conceive in abstract systems in FSMs. Determinism for CPS applications loosely translates to the requirement that a fixed number of events (e.g., fixed number message arrival per cycle) obey the given timing constraints, i.e., they happen within their respective time windows (not early or late). However, real world systems cannot offer absolute certainty on the adherence of such timing constraints. In our view, deterministic communication systems ensure that the likelihood of successful message transmission between different CPS application endpoints within their respective/expected time windows is at least equal to the application requirements. This necessitates that uncertainty in the system can be bounded over some time horizon rendering these systems *predictable*. For example, an Ethernet cable is predictable as it is possible to bound packet error rate of an Ethernet cable for a given Signal-to-Noise ratio.

7.2 Determinism in 5G

Traditionally, wired communication systems such as proprietary fieldbus systems, e.g., Powerlink, EtherCat, etc., and standardized Ethernet systems, e.g., TSN, have been chosen for deterministic communications, as their stochastic fluctuations can be engineered towards application requirements. However, this is not the case with wireless systems where spatial-temporal random factors such as interference, mobility, fading, and shadowing (leading to variable signal attenuation) subject the communication system to much higher stochastic uncertainty.

As a starting point for the investigation of 6G architecture principles towards deterministic communication, it is valuable to review the architecture design of the 5G system for deterministic communication. One important component of 5G for deterministic communication is Ultra Reliable and Low Latency Communications (URLLC). URLLC is primarily providing functionality in the radio access network to increase diversity and provide extra robustness for transmission of control and data information over the radio channel, in combination with features that reduce latency. However, the URLLC functionality for a radio link has limited direct impact on the system architecture. From an architecture perspective, design principles for deterministic communication were driven by adding the support for TSN into 5G (see Figure 7.1**Fehler! Verweisquelle konnte nicht gefunden werden.**). The principles of support of TSN have been extended for any time-sensitive communication (TSC) service (not only being based on Ethernet as TSN), and also the explicit support of DetNet has been added. In addition, means for redundant communication paths were introduced in 5G to provide high availability and resilience. The general principles that are applied are the following:

- Traffic classification and QoS [5GA21c], [5GS20-D51], [5GS20-D52], [5GS21-D54], [3GPPY23-23501]:
    - Traffic can be separated at the 5G ingress via traffic flow filters into so-called QoS flows, which can be specified via e.g., network exposure APIs. Specific service (and performance) requirements can thus be defined for different traffic flows.
    - Traffic treatment in the 5GS is adapted to the needs on a per-QoS-flow level, which allows for configuration of the 5G network functions in a service-specific way.
    - Resource management distributes resources to the QoS flows according to the specified needs.
    - Redundant communication paths can be provided through the 5G system over either RAN or the entire 5G network.
- Interaction of an (industrial) application framework with the network via network exposure [SSD+23], [KDS+23], [5GA21a], [5GS20-D52], [5GS21-D55], [3GPP23-23501], [3GPP23-23434]:
    - An application framework that is used end-to-end for the handling of application services, can configure the connectivity of the devices over 5G, including the QoS configuration, but also functionality like time synchronization.
    - The exposure interface also allows to provide network information to the application framework.
- Support for interworking with deterministic communication frameworks [5GS20-D51], [5GS20-D52], [5GS21-D53], [5GS21-D55], [5GA21b], [3GPP23-23501]:
    - Most deterministic communication frameworks are today built on Ethernet, and TSN is foreseen as the future converged Ethernet standard enabling time-sensitive deterministic communication.
    - 5G adopted interworking with TSN to support transmission of Ethernet traffic, perform Ethernet bridging and allow traffic configuration according to the TSN standard – providing a TSN specific “network exposure” where the 5G capabilities are reported and traffic flows are configured according to the TSN standard.
    - The mechanisms of TSN support have been extended to address IP based DetNet communication to interwork from 5G with the traffic handling mechanisms defined in IETF DetNet.
    - For TSN and DetNet the interaction of the 5G system with the TSN / DetNet network controller (i.e., TSN CNC and DetNet controller) is extended to comply with the network control and configuration frameworks used by TSN (between CNC and TSN bridge) and DetNet (between DetNet Controller and DetNet node)
- Support for time synchronization [5GS20-D51], [5GS20-D52], [PDR+21], [MAG+19], [GLR+20], [3GPP23-23501]:

  - Time synchronization is a common functionality used in deterministic applications for synchronizing application endpoints to a common time reference, but also to synchronize communication nodes (e.g., TSN bridges) to a common time for network configuration and control.
  - Passing (g)PTP through the 5G system.
  - Using 5G as (PTP) time reference to external system.
  - Extending the time-synchronized RAN to 5GS user plane (UE, UPF). Time error control.
- Non-public networks [5GS20-D52], [3GPP23-23501], [5GA21d]:
  - Many use cases for deterministic communication appear in local confined areas within closed user groups that belong to one entity. Examples are smart factories, ports, or mines, where a private 5G network deployment is desired. To enable this, non-public networks (NPN) have been introduced in 5G, which allow to provide private networking services to a predefined group of devices and connected to a private system. Two variants of NPN have been specified, a standalone NPN (SNPN) as separate private network deployment, and a public-network integrated NPN (PNI-NPN), which is realized by providing a private network service that at least partly uses a public network infrastructure and is separated via network slicing.

All of these principles are realized in 5G Stand Alone (5G SA), whereas in 5G Non-Stand-Alone (5G NSA) – i.e., the majority of 5G networks – not all principles might be available.

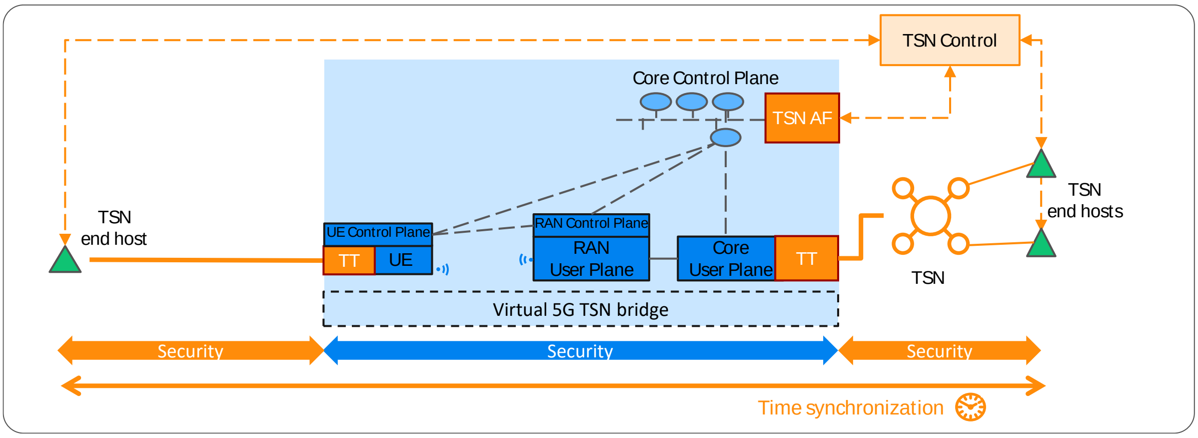


Figure 7.1: 5G deterministic communication for end-to-end TSN services.

On one hand, the proposed architectural enhancements in 5G add support for the integration with other deterministic communications technologies (e.g., TSN and DetNet), on the other hand, URLLC functionalities provide margins towards stochastic variations in the wireless segments. However, this comes at a cost of lower resource efficiency: reduction of stochastic variance leads to a reduction of the capacity (e.g., in terms of the number of applications being served) that can be provided by the wireless system. Though 5G URLLC in general enables more reliable wireless communications (for given conditions) in contrast to previous generations of cellular networks, the Packet Delay Variation (PDV) of 5G systems is still orders of magnitude higher as compared to wired networks. Overall, the performance of 5G systems is highly dynamic in space and time due to the dependency on the spatio-temporal propagation environment and the resulting stochastic nature of wireless communications. This contrasts with wired communication systems where performance is mostly isolated from local conditions.

7.3 Guiding principles for 6G deterministic communications

While 5G represents a significant improvement towards deterministic applications – in comparison to previous cellular network generations – even 5G URLLC systems have substantial stochastic variations. This is a fundamental aspect of wireless systems that cannot be overcome. Consequently, periods with higher and lower reliability (for example for a given target latency) will occur. A more efficient approach might hence be to allow adaptation among applications towards communication service variability. Nevertheless, a prerequisite for this adaptation process is the necessity to anticipate communication service variability to some degree at run-time, leveraging the specific spatio-temporal characteristics of a given deployment. The approach taken by the DETERMINISTIC6G project is to devise architectures and algorithms that enable this predictability, and subsequently to expand towards adaptation of deterministic and other applications. In the following, we detail the aspect of predictability, the spectrum of applications, and how to leverage predictability in the context of the different application classes.

At a high level, predictability of communication systems refers *to the ability to accurately predict parameters of stochastically evolving KPIs of a given communication system over a given time interval at run-time*. For example, parameters of interest might be the average throughput for a given terminal over a certain time interval, the likelihood of subsequent correct packet reception over a certain time interval or the latency tail probability of an instantaneous packet transmission. Accuracy of a parameter prediction must be measured against ground-truth, leading to a maximal-acceptable average error between the predicted parameters and its ground-truth value. Important is furthermore the dependency between design-time and run-time: On the one hand, the predictability of any communication system is dependent on the design of the system (i.e., the specification of a communication protocol), and on the other hand, on the algorithms and information base leveraged at run-time. This is particularly relevant with respect to the upcoming 6G standardization.

Furthermore, predictability is application dependent. 6G systems will support many different application types where the requirements of these applications form a wide spectrum and can be time varying. Here, the focus is on the industrial domain, since it already combines many of those challenges, which target deterministic E2E communication in converged wireless and wired networks. However, it is expected that also many other fields of application will benefit from advancements in the 6G architecture and new features that lead to improved E2E determinism of communication. At a higher level, three different traffic classes have been defined in 5G for the industrial domain: Deterministic periodic, deterministic non-periodic and non-deterministic, depending on the type of control and the need for determinism [3GPP23-22104]. In addition, integrated communication & compute applications like exoskeletons or extended reality applications are emerging in the industrial context, which will certainly also be relevant for 6G systems.

In a complex and/or dynamic use case environment – as in the focus of this project – interactions between applications, but also between an application and the use case environment (especially the communication system) itself, may change over time. For instance, adding or removing entities to the system will also come with implications on the usage of communication resources. Due to restricted resources (e.g., communication bandwidth, CPU time), physical limitations (e.g., signal propagation characteristics), component availabilities, and faults, some of the applications may fail to continue providing their service. To prevent this, a management entity in the system should be able to monitor and predict system behaviour at runtime in order to perform corrective measures (e.g., adapting resource allocation / schedules) before the appearance of an undesired situation.

As an example, an application that performs obstacle detection on a video stream might specify the requirement for a high-quality frame rate of 50 frames per second. Alternatively, it supports a low-quality mode which only requires 25 frames per second. However, the consequence in this example would be that in the lower-quality mode the worst-case detection time for obstacles is longer. Thus, an application that uses information about detected obstacles also needs to operate at a lower performance mode (e.g., lower maximum speed of an AGV). Based on the prediction of the bandwidth that can be deterministically delivered (i.e., probability of packet loss below a certain value), the management entity may select one of these QoS modes and forward the decision to the application.

However, the nature and degree of predictability required for all these application types differs. For instance, a motion control application might require precise forecasting in terms of latency over an extended duration at the packet level, while an obstacle detection application might be more tolerant to delay fluctuations every now and then thus might only require an accurate prediction regarding the average latency regardless of individual packet delays. Both applications have furthermore potentially different prediction time horizons of interest. In other words, this kind of system management is only possible and efficient if the detailed requirements of each individual application are known. Only when these requirements can be fulfilled, for instance by some underlying services, then the application can operate correctly and serve its purpose.

When analysing the above-mentioned use cases, one can observe a high possibility for dynamic changes of the entities in the system and the corresponding communication relations. Therefore, a standardized description of services is necessary that allows an application to register its requirements at the network management.

7.4 Architecture aspects of 6G deterministic communications

The principles and functionality introduced by 5G presented previously lay the foundation for reliable wireless communication and shall be evolved for 6G to address better deterministic communication needs.

It is desired that an application can depend on the 6G system as a communication and compute platform to achieve its requirements for determinism. As explained above, this has the following implications:

- The application must be able to describe its requirements on reliability,
- the 6GS must be deterministic in order to be able to prove that it fulfils the requirements on reliability.

To this end, some new capabilities must be provided for 6G. Creating programmable 6G systems to allow deterministic services to express their requirements and preferences. Deterministic applications are developing increasingly to larger flexibility, including dynamic adaptations of operation modes and requirements. The interface towards the 6G system must be sufficiently rich to express the trade-offs the applications can tolerate in its operation and allow for dynamic adaptations. Network exposure shall provide such functionality in a sufficiently abstract and simple form, to allow, e.g., configuration of network and compute capabilities (extending approaches in [SSD+23] [KDS+23]. In some cases, deterministic applications may be provided by dedicated 6G network deployments (e.g., SNPN). But such services may also be provided by using and configuring 6G network platforms provided by communication service providers, e.g., via PNI-NPN. In this case, the programmability of the network platform should include business support capabilities to enable required business agreements among involved parties (see e.g., [FMK23]).

A further requirement to provide dependable communication and compute to deterministic applications, is to be able to predict and prove that the deterministic service requirements can be met by the 6G network platform. This calls for significant improvements compared to 5G systems. On the one hand, observability of the performance delivered to deterministic services needs to be provided via e.g., network and QoS monitoring. But on the other hand, to be able to prove that performance bounds can be met, a prediction of the achievable QoS for a certain time horizon is needed. For this, machine-learning provides a feasible foundation where collected observations of network performance and states, potentially enhanced with further environmental awareness and contextual information, are used as basis for prediction of the communication and compute service for the next time period. A peculiarity of deterministic services is that such prediction must be able to include tail probabilities due to the high demand on reliability and availability. A data-driven architecture design [Roe20] is needed for collecting network probes, applying ML-based analysis and prediction. Constraints on exchange of distributed data should be considered to achieve a good trade-off of centralized and distributed AI. Due to the time-critical nature of deterministic services, time awareness is a central capability, not only for deterministic service provisioning but also for the observability of network characteristics. Time synchronization, spanning over the network domain, as well as the application domain, is a foundation for such time awareness and can build on the principles established with 5G.

The attractive capabilities of scalable and cost-efficient compute via cloud-computing approaches are being increasingly adopted in the application domain and in cloud-based networking. Furthermore, distributed cloud infrastructures are being integrated with network infrastructures. 6G is expected to be based on a cloud-native design and applications are increasingly provided in edge cloud platforms. Principles for providing time-aware deterministic services to cloud-hosted applications and connecting them via deterministic networks is an area for further investigation (see e.g. [5GS20-D52], [5GS21-D54]). Due to requirements on high reliability and availability, the 6G architecture shall provide capabilities for network reliability, availability, and resilience (NRAR) [BSL+20]. This leads to design objectives of avoiding single-points-of-failures and minimizing dependencies on functional operation, e.g., between the control plane and the user plane functions of the network. This includes a principle of security-by-design to monitor, analyse, prohibit, and mitigate security threats.

7.5 Challenges

Given all these characteristics and dependencies, the central challenge arising from our approach is how to leverage predictability. For instance, initial predictability at run-time towards certain KPIs from applications can help the 6G system to manage resources to sustain application performance. For instance, instantaneous availability might be predicted for a given system. Assuming this prediction to be accurate over a relevant time horizon, the communication system could leverage the prediction to allocate more resource blocks or change other scheduling parameters in order to improve or sustain availability. Beyond communication system management based on predictability, a further opportunity is to determine through a dynamic application interface new operational requirements to fulfil, for instance in the case that given requirements are likely not going to be met according to run-time predictions. In the application spectrum, different applications are more flexible, while others are less flexible. Nevertheless, it is a distinct target of DETERMINISTIC6G to identify for all relevant industrial applications foreseeable today, how the adaptation between system performance and application requirement could be structured. Finally, it should be noted that such a flexible management also comprises of the possibility to tighten application requirements if communication system conditions allow for this.

In the course of the project, the value of individual KPIs, communication parameters, and system capabilities for improving the predictions about the behaviour (i.e., expected packet delay or packet losses) of the communication system, and for performing appropriate reactions by the management entity, will be evaluated. Based on this evaluation, the most valuable set of KPIs, parameters, and capabilities will be used to define a novel service description for deterministic communication. Additionally, the feasibility of specifying multiple possible QoS levels at the same time will be elaborated. In such a scenario, an application could provide a service request containing multiple combinations of the parameters in the service definition. Each parameter set refers to a QoS level which is supported by the application. The management entity within the communication system may then decide which QoS level can be supported with a certain probability. This allows that an application may still operate in a degraded (e.g., less efficient) mode instead of completely ceasing operation. Dynamic change-over of operation modes at runtime needs to be supported by the application itself and must be performed within a specified lead time.

Also, the new concept of KVIs introduces the possibility to consider the social impact of an application in the description of services. During the project it will be evaluated if it is feasible to bring certain social KVs into a relation with service requirements of an application that is part of a more complex use case domain (e.g., smart farming). The corresponding KVIs may then be integrated into the service description. This could lead, for instance, to a higher prioritization of more valuable applications in a use case, if the corresponding KVIs indicate a gain in the social value. In contrast, losses in KVIs would then result in lower priorities or even rejection of the application by the network management entity.

8 Conclusions

Within this document, four visionary use cases have been presented, which focus on the assistance of humans in physically challenging tasks, the collaboration and interaction of humans with the next generation of digitalized production systems, and the extended mobile and dynamic behaviour of future industrial automation facilities. These use cases highlight the need for a deterministic E2E communication infrastructure, which is only possible with technological enhancements in existing wireless and wired communication standards to support improved predictability and to initiate appropriate reactions, if the anticipated QoS changes significantly. Further key elements enabling these use cases are the convergence of 6G wireless communication with wired network infrastructure, the consideration of cloud computing in the E2E communication scenarios, and the provision of corresponding security by design.

Due to the early stage of the research towards deterministic E2E communication in 6G networks, the illustrated use cases may only provide an abstract view on the opportunities created, once the corresponding technology is available. However, they are already suitable to drive the search for technical solutions and the investigation of their social impact. As the DETERMINISTIC6G project progresses, deeper insights into specific aspects of the use cases will be necessary and corresponding elaborations will follow.

In addition to the use case descriptions, the document summarizes the challenges that need to be tackled in order to achieve the specified goal of a deterministic E2E communication infrastructure. It is further outlined, which topics are in the focus of research that is carried out and which architectural concepts will be developed throughout the project. Based on the use cases, KVIs are defined that lead to a better understanding of the social impact of the involved technologies, and thus, allow to provide guidelines for boosting positive effects and preventing negative consequences.

The specification of a deterministic service description facilitates the dynamic adaptation of the system configuration to changing requirements, priorities, and system capabilities. Novel architectural concepts are introduced to achieve better predictability of the QoS and to provide new means for letting the system automatically react to changing situations. Since the extension of system features may come with additional threats, a thorough assessment of cyber security requirements and possible threat mitigation strategies is inevitable and is proposed to be conducted as part of the project.

Appendix

References

[3GPP23-22104] 3GPP, "Service requirements for cyber-physical control applications in vertical domains", technical specification, TS 22.104, March 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3528

[3GPP23-23434] 3GPP TS 23.434, "Service Enabler Architecture Layer for Verticals (SEAL); Functional architecture and information flows", April 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144

[3GPP23-23501] 3GPP TS 23.501, " System architecture for the 5G System (5GS)", technical specification, April 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144

[3GPP23-23502] 3GPP TS 23.502, "Procedures for the 5G System (5GS)", technical specification, April 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145

[3GPP23-23548] 3GPP TS 23.548, "5G System Enhancements for Edge Computing; Stage 2", technical specification, April 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3856

[3GPP23-23558] 3GPP TS 23.548, "Architecture for enabling Edge Applications", technical specification, March 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3723

[3GPP28-23538] 3GPP TS 28.538, "Management and orchestration; Edge Computing Management", technical specification, March 2023, https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3910

[5GA21a] 5G-ACIA, "Exposure of 5G Capabilities for Connected Industries and Automation Applications", white paper, February 2021, https://5g-acia.org/whitepapers/exposure-of-5g-capabilities-for-connected-industries-and-automation-applications-2/

[5GA21b] 5G-ACIA, "Integration of 5G with Time-Sensitive Networking for Industrial Communications", white paper, February 2021, https://5g-acia.org/whitepapers/exposure-of-5g-capabilities-for-connected-industries-and-automation-applications-2/

[5GA21c] 5G-ACIA, "5G QoS for Industrial Automation", white paper, November 2021, https://5g-acia.org/whitepapers/5g-quality-of-service-for-industrial-automation-2/

[5GA21d] 5G-ACIA, "5G Non-Public Networks for Industrial Scenarios", white paper, September 2021, https://5g-acia.org/resources/whitepapers-deliverables/

[5GS20-D51] 5G-SMART Deliverable 5.1, "First report on new technological features to be supported by 5g standardization and their Implementation impact", May 2020, https://5gsmart.eu/deliverables/

[5GS20-D52] 5G-SMART Deliverable 5.2, "First Report on 5g network architecture options and assessments", November 2020, https://5gsmart.eu/deliverables/

[5GS21-D53] 5G-SMART Deliverable 5.3, "Second report on new technological features to be supported by 5g standardization", November 2021, https://5gsmart.eu/deliverables/

[5GS21-D54] 5G-SMART Deliverable 5.4, "Second report on 5g network architecture options and assessments", November 2021, https://5gsmart.eu/deliverables/

[5GS21-D55] 5G-SMART Deliverable 5.5, "Report describing the framework for 5G system and network management functions", November 2021, https://5gsmart.eu/deliverables/

[6GIA22] 6G Infrastructure Association, "What societal values will 6G address?", whitepaper, May 2022, https://5g-ppp.eu/wp-content/uploads/2022/05/What-societal-values-will-6G-address-White-Paper-v1.0-final.pdf

[AA17] M. Abraham, M. Annunziata, "Augmented Reality Is Already Improving Worker Performance", Harvard Business Review, March 2017, https://hbr.org/2017/03/augmented-reality-is-already-improving-worker-performance

[AA22] Agnius Birutis, Anders Mykkeltveit, "Practical Jamming of a Commercial 5G Radio System at 3.6 GHz", International Conference on Military Communications and Information Systems (ICMCIS 2022), September 2022, https://doi.org/10.1016/j.procs.2022.09.007

[AKP+21] F. Alriksson, D.H. Kang, C. Phillips, J.L. Pradas, A. Zaidi, "XR and 5G: Extended reality at scale with time-critical communication", Ericsson Technology Review, August 2021, https://www.ericsson.com/4a492d/assets/local/reports-papers/ericsson-technology-review/docs/2021/xr-and-5g-extended-reality-at-scale-with-time-critical-communication.pdf

[ADF+23] F. Alriksson, E. Drugge, A. Furuskär, D.H. Kang, J. Kronander, J.L.Pradas, Y.Sun, "Future network requirements for extended reality use cases", Ericsson Technology review, April 2023, https://www.ericsson.com/496150/assets/local/reports-papers/ericsson-technology-review/docs/2023/future-network-requirements-for-xr-apps.pdf

[BSL+20] K. Balachandran, B. Smeets, M. Liljenstam, J. Ylitalo, E. Fogelström, "Building trustworthiness into future mobile networks", white paper, July 2020, https://www.ericsson.com/en/reports-and-papers/white-papers/building-trustworthiness-into-future-mobile-networks

[Bus23] Buscemi, A., Ponaka, M., Fotouhi, M., Koebel, C., Jomrich, F. and Engel, T., "An Intrusion Detection System Against Rogue Master Attacks on gPTP", IEEE Vehicular Technology Conference (VTC2023-Spring), June 2023, http://hdl.handle.net/10993/54718

[Cre21] Crea, Simona, et al. "Occupational exoskeletons: A roadmap toward large-scale adoption. Methodology and challenges of bringing exoskeletons to workplaces." Wearable Technologies, Volume 2 (2021). https://doi.org/10.1017/wtc.2021.11

[dLBK+16] de Looze MP, Bosch T, Krause F, Stadler KS and O'Sullivan LW (2016), "Exoskeletons for industrial application and their potential effects on physical work load", Ergonomics 59, 671–681., October 2015, http://dx.doi.org/10.1080/00140139.2015.1081988

[ECL+22] Ericsson Consumer Lab study, "5G: The next wave", October 2022, https://www.ericsson.com/en/reports-and-papers/consumerlab/reports/5g-next-wave

[EU2023] European Commission, "A New Industrial Strategy for Europe", March 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0102 .

[EWCS15] Sixth European Working Condition Survey https://www.eurofound.europa.eu/surveys/european-working-conditions-surveys/sixth-european-working-conditions-survey-2015

[FIS21] K. Fischer, A. Furch, L. Lindemann, O. Pfaff, T. Pössler , G. Steindl, "Security for IEC/IEEE 60802 Overview of Approach", presentation in IEEE 802.1 Working Group, March 2021, https://grouper.ieee.org/groups/802/1/files/public/docs2021/60802-Pfaff-et-al-Security-Overview-0321-v01.pdf

[FMK23] J. Friman, E. Mueller and B. van Kaathoven, "Monetizing API exposure for enterprises with evolved BSS," in *Ericsson Technology Review*, vol. 2023, no. 1, pp. 2-10, January 2023, doi: 10.23919/ETR.2023.10035875. https://ieeexplore.ieee.org/document/10035875

[GD08] Grieve JR and Dickerson CR, "Overhead work: Identification of evidence-based exposure guidelines", Occupational Ergonomics 8(1), 53–66, January 2008, https://doi.org/10.3233/OER-2008-8105

[GLR+20] I. Godor, M. Luvisotto, S. Ruffini, K. Wang, D. Patel, J. Sachs, O. Dobrijevic, D. P. Venmani, O. Le Moult, J. Costa-Requena, A. Poutanen, C. Marshall, J. Farkas, "A Look Inside 5G Standards to Support Time Synchronization for Smart Manufacturing," in *IEEE Communications Standards Magazine*, vol. 4, no. 3, pp. 14-21, September 2020, doi: 10.1109/MCOMSTD.001.2000010. https://ieeexplore.ieee.org/document/9204594

[GTP+20] Grazi L, Trigili E, Proface G, Giovacchini F, Crea S and Vitiello N, "Design and experimental evaluation of a semi-passive upper-limb exoskeleton for workers with motorized tuning of assistance", IEEE Transactions on Neural Systems and Rehabilitation Engineering 28, 2276–2285, October 2020, http://dx.doi.org/10.1109/tnsre.2020.3014408

[Gya22] Gyamfi, E. and Jurcut, A., "Intrusion Detection in Internet of Things Systems: A Review on Design Approaches Leveraging Multi-Access Edge Computing", Machine Learning, and Datasets. *Sensors*, *22*(10), p.3744., May 2022, https://doi.org/10.3390/s22103744

[HML+20] Howard J, Murashov VV, Lowe BD and Lu ML, “Industrial exoskeletons: Need for intervention effectiveness research”, American Journal of Industrial Medicine 63(3), 201–208., December 2019, https://doi.org/10.1002/ajim.23080

[Hoe03] C. Hoefer, “Causal Determinism”, January 2003; https://plato.stanford.edu/entries/determinism-causal/

[KDS+23] A. Karaagac, O. Dobrijevic, D. Schulz, G. Seres, A. Nazari, H. Przybysz, J. Sachs, "Managing 5G Non-Public Networks from Industrial Automation Systems”, 2023 IEEE 19th International Conference on Factory Communication Systems (WFCS), Pavia, Italy, 2023, pp. 1-8, doi: 10.1109/WFCS57264.2023.10144248. https://ieeexplore.ieee.org/document/10144248

[LÖV23] Lövehagen, Nina, A framework for mapping potential sustainability impact of digitalization solutions, June 2023, https://ssrn.com/abstract=4486120

[LSW+22] J. Leng, W. Sha, B. Wang, P. Zheng, C. Zhuang, Q. Liu, T. Wuest, D. Mourtzis, L. Wang, „Industry 5.0: Prospect and retrospect”, Journal of Manufacturing Systems, vol. 65, pp. 279-295, October 2022, doi: https://doi.org/10.1016/j.jmsy.2022.09.017

[MAD20] Monica L, Anastasi S and Draicchio F, “Occupational exoskeletons: Wearable robotic devices and preventing work-related musculoskeletal disorders in the workplace of the future”, pp. 1–12., September 2020, https://osha.europa.eu/en/publications/occupational-exoskeletons-wearable-robotic-devices-and-preventing-work-related

[MAG+19] A. Mahmood, M. I. Ashraf, M. Gidlund, J. Torsner and J. Sachs, "Time Synchronization in 5G Wireless Edge: Requirements and Solutions for Critical-MTC," in IEEE Communications Magazine, vol. 57, no. 12, pp. 45-51, December 2019, doi: 10.1109/MCOM.001.1900379. https://ieeexplore.ieee.org/document/8930825

[Mar23] W. Martino, “Adaptive vs Flexible manufacturing”, June 2023, https://www.br-automation.com/en/news/blog/what-do-you-know-about-adaptive-manufacturing/

[MLT+22] F. Missiroli et al., "Rigid, Soft, Passive, and Active: A Hybrid Occupational Exoskeleton for Bimanual Multijoint Assistance," in IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 2557-2564, April 2022, doi: 10.1109/LRA.2022.3142447.

[Pat23] L. Patsakas, “Impact of adaptive manufacturing on the field of medical device assembly”, 14.06.2023, https://www.br-automation.com/en/news/blog/we-doubled-their-output-from-the-same-footprint/

[Par+17] A. Parent-Thirion, et al., “Sixth European Working Conditions Survey – Overview report,” Eurofound, 2017, Ireland, doi: https://doi.org/10.2806/422172

[PDR+21] D. Patel, J. Diachina, S. Ruffini, M. De Andrade, J. Sachs and D. P. Venmani, "Time error analysis of 5G time synchronization solutions for time aware industrial networks”, 2021 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control, and Communication (ISPCS), NA, FL, USA, 2021, pp. 1-6, doi: 10.1109/ISPCS49990.2021.9615318. https://ieeexplore.ieee.org/document/9615318

[Res20] Research Park University of Illinois Urbana Champaign, EarthSence TerraSentia Features in Successful Farming, July 2020, https://researchpark.illinois.edu/article/earthsense-terrasentia-featured-in-successful-farming/

[Roe20] D. Roeland, "An introduction to data-driven network architecture", October 2020, https://www.ericsson.com/en/blog/2020/10/data-driven-network-architecture

[Sha23] Aljoharah Al Shaikh, Pure Harvest smart farms thrive with ABB Ability Energy Manager, 25.01.2023, https://new.abb.com/news/detail/98882/pure-harvest-smart-farms-thrive-with-abb-ability-energy-manager

[Sha23b] Sharma, G.P., Patel, D., Sachs, J., De Andrade, M., Farkas, J., Harmatos, J., Varga, B., Bernhard, H.P., Muzaffar, R., Atiq, M.K. and Duerr, F., Towards Deterministic Communications in 6G Networks: State of the Art, Open Challenges and the Way Forward, April 2023, *arXiv preprint arXiv:2304.01299*.

[SSD+23] G. Seres, D. Schulz, O. Dobrijevic, A- Karaağaç, H. Przybysz, A. Nazari, P. Chen, M. L. Mikecz, A. D. Szabó, " Creating programmable 5G systems for the Industrial IoT", in *Ericsson Technology Review*, vol. 2022, no. 10, pp. 2-12, October 2022, doi: 10.23919/ETR.2022.9934828. https://ieeexplore.ieee.org/document/9934828

[Wav16] J.M.P. van Waveren, The Asynchronous Time Warp for Virtual Reality on Consumer Hardware, November 2016, https://dl.acm.org/doi/pdf/10.1145/2993369.2993375

List of abbreviations

AGV	Automated Guided Vehicle
AI	Artificial Intelligence
AR	Augmented Reality
AV	Aerial Vehicle
CNC	Centralized Network Controller
CPS	Cyber-Physical System
CUC	Centralized User Configuration
DL	Downlink
DoS	Denial of Service
E2E	End-to-End
ERP	Enterprise Resource Planning
EV	Electric Vehicle
FPGA	Field Programmable Gate Array
FSM	Finite State Machine
FX	Field eXchange
GPP	General Purpose Processor
GPU	Graphics Processing Unit
GV	Ground Vehicle
HMD	Head Mounted Device
HVAC	Heating Ventilation and Air Conditioning
ICT	Information and Communication Technology
IP	Internet Protocol
KPI	Key Performance Indicator
KV	Key Value
KVI	Key Societal Value Indicator
LIDAR	Light Detection and Ranging
MEC	Multi-Access Edge Computing
ML	Machine Learning
MPM	Mobile Processing Module
MTBF	Mean Time Between Failures
MTP	Motion-to-Photon
NPN	Non-public Networks
NSA	Non-Stand-Alone
OE	Occupational Exoskeleton
OPC UA	Open Platform Communications Unified Architecture
OS	Operating System
PDV	Packet Delay Variation
PNI-NPN	Public network integrated NPN
QoE	Quality of Experience

QoS	Quality of Service
RAN	Radio Access Network
SA	Stand Alone
SDG	Sustainability Development Goal
SNPM	Standalone NPM
SLA	Service Level Agreement
SSLA	Security Service Level Agreement
TRL	Technology Readiness Level
TSC	Time-Sensitive Communication
TSN	Time-Sensitive Networking
UAV	Unmanned Aerial Vehicle
UCS	Use Case Scenario
UGV	Unmanned Ground Vehicle
UL	Uplink
URLLC	Ultra Reliable and Low Latency Communications
VM	Virtual Machine
VNF	Virtual Network Function
VR	Virtual Reality
VRT	Variable Rate Treatment
WRMD	Work-related Musculoskeletal Disorders
XR	Extended Reality