ZEROSWARM

ZERO-enabling Smart networked control framework for Agile cyber physical production systems of systems

D3.1 Enhancing 5G NPN communications as CPSoS middleware

Project funded by Horizon Europe, Grant Agreement #101057083

ZEROSWARM

Topic	HORIZON-CL4-2021-TWIN-TRANSITION-01-08
Project Title	ZERO-enabling Smart networked control framework for Agile cyber physical production systems of systems
Project Number	101057083
Project Acronym	Zero-SWARM
Contractual Delivery Date	M14
Actual Delivery Date	M14
Contributing WP	WP3 – Shaping 5G in the Industrial Domain
Project Start Date	01/06/2022
Project Duration	30 Months
Dissemination Level	Public
Nature of Deliverable	Report

Document History

Date	*Version*	*Author*	*Description*
04-01-2023	*0.1*	*Huawei & all*	*ToC definition*
09-03-2023	*0.2*	*Huawei & all*	*Collaboration distribution*
30-04-2023	*0.3*	*i2CAT*	*Contribution to Section 4*
11-05-2023	*0.4*	*Huawei*	*Contribution to Section 2*
12-05-2023	*0.5*	*Huawei, Neutroon*	*Contribution to Section 3.1*
14-06-2023	*0.6*	*UMH*	*Contribution to Section 3.2*
27-06-2023	*0.7*	*Huawei*	*Contribution to Section 3.3*
03-07-2023	*0.8*	*LTU*	*Contribution to Section 3.3*
12-07-2023	*0.9*	*Huawei, UMH, Neutroon*	*Updates of Sections 2 and 3*
13-07-2023	*1.0*	*Huawei*	*Introduction*
17-07-2023	*1.1*	*Huawei*	*Conclusion*
24-07-2023	*1.2*	*Huawei*	*Editing whole Deliverable*
27-07-2023	*1.3*	*i2CAT, CERTH*	*Internal review of Deliverable*
28-07-2023	*1.4*	*Huawei*	*Addressing comments of Internal reviewers*
28-07-2023	*2.0*	*CERTH*	*Final submission*

Project funded by Horizon Europe, Grant Agreement #101057083

Author List

Leading Author (Editor)			
Surname	Initials	Beneficiary Name	Contact email
Krendzel	AK	Huawei	andrey.krendzel@huawei.com
Co-authors (in alphabetic order)			
Surname	Initials	Beneficiary Name	Contact email
Khodashenas	PSK	Huawei	pouria.khodashenas@huawei.com
Contributors (in alphabetic order)			
Surname	Initials	Beneficiary Name	Contact email
Camps	DC	i2CAT	daniel.camps@i2cat.net
Guerra-Gomez	RGG	Neutroon	rolando.guerra@neutroon.com
Gozalvez	JG	UMH	j.gozalvez@umh.es
Khodashenas	PSK	Huawei	pouria.khodashenas@huawei.com
Krendzel	AK	Huawei	andrey.krendzel@huawei.com
Liakh	LT	LTU	tatiana.liakh@ltu.se
Lucas-Estañ	MCLE	UMH	m.lucas@umh.es
Mendes	BM	Ubiwhere	bmendes@ubiwhere.com
Rigazzi	GR	Huawei	giovanni.rigazzi@huawei.com
Sepulcre	MS	UMH	msepulcre@umh.es
Zhuqianghua	ZQ	Huawei	zhuqianghua@huawei.com

Reviewers List

List of reviewers (in alphabetic order)			
Surname	Initials	Beneficiary Name	Contact email
Carmona	EC	i2CAT	estela.carmona@i2cat.net
Mpatziakas	AM	CERTH	ampatziakas@iti.gr

DISCLAIMER OF WARRANTIES

This document has been prepared by Zero-SWARM project partners as an account of work carried out within the framework of the contract no 101057083.

Neither Project Coordinator, nor any signatory party of Zero-SWARM Project Consortium Agreement, nor any person acting on behalf of any of them:

- makes any warranty or representation whatsoever, express or implied,
    - with respect to the use of any information, apparatus, method, process, or similar item disclosed in this document, including merchantability and fitness for a particular purpose, or
    - that such use does not infringe on or interfere with privately owned rights, including any party's intellectual property, or
- that this document is suitable to any particular user's circumstance; or
- assumes responsibility for any damages or other liability whatsoever (including any consequential damages, even if Project Coordinator or any representative of a signatory party of the Zero-SWARM Project Consortium Agreement, has been advised of the possibility of such damages) resulting from your selection or use of this document or any information, apparatus, method, process, or similar item disclosed in this document.

Zero-SWARM has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101057083. The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the deliverable lies entirely with the author(s).

ZEROSWARM

Table of Contents

Project funded by Horizon Europe, Grant Agreement #101057083

List of Abbreviation

3GPP	3rd Generation Partnership Project
5G	5th Generation of mobile system
5GC	5G Core Network
5GS	5G System
5G-ACIA	5G Alliance for Connected Industries and Automation
5QI	5G QoS Identifier
AF	Application Function
AM	Additive Manufacturing
API	Application Programming Interface
AR	Application Relation
BAT	Burst Arrival Time
CC	Conformance Class
CG	Configured Grant
CIP	Common Industrial Protocol
CR	Communication Relation
CNC	Central Network Configuration
CPSoS	Cyber-Physical System of Systems
CRC	Cyclic Redundancy Check
CSMA/CD	Carrier Sensing Multiple Access and Collision Detection
CUC	Centralized User Configuration
DC	Distributed Clock
DCS	Distributed Control Systems
DL	Downlink
DN	Data Network
DS-TT	Device side TSN Translator
DSCP	Differentiated Services CodePoint (DSCP)
ECC	Execution Control Chat
ED	End Devices
ESC	EtherCAT Slave Controller
eMBB	Enhanced Mobile Broadband
ETSI	European Telecommunications Standards Institute
FB	Functional Block
FW	Firewall
gNB	new generation Node B
GSD	General Station Descriptor
GTP	GPRS Tunnelling Protocol
HP	HyperPeriod
ICT	Information and Communications Technology
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IIoT	Industrial Internet of Things
IRT	Isochronous Real-Time
LAN	Local Area Network
LBO	Local Breakout
LLDP	Link Layer Discovery Protocol
MCS	Modulation and Coding Scheme
MDBV	Maximum Data Burst Volume
mMTC	Massive Machine-Type Communications
MNO	Mobile Network Operator
MRP	Media Redundancy Protocol (MRP)
NAT	Network Address Translation
NEF	Network Exposure Function
NG	Next Generation RAN

NPN	Non-Public Network
NR	New Radio
NRT	Non-Real Time
NS-TT	Network side TSN Translator
NW	Network
ODVA	Open Device Vendors Association
OFDM	one Orthogonal Frequency Division Multiplexing
OP	Offloading Point
OPC	Open Platform Communications
OPC UA	OPC United Architecture
OT	Operational Technology
P4	Programming Protocol-independent Packet Processors
PCF	Policy Control Function
PDB	Packet Delay Budget
PDCP	Packet Data Convergence Protocol
PDU	Packet Data Unit
PER	Packet Error Rate
PID	Proportional Integral Derivative
PLC	Programmable Logic Controller
PNI-NPN	Public Network Integrated Non-Public Network
PSA	PDU Session Anchor
PTP	Precision Time Protocol
QoD	Quality on Demand
QoS	Quality of Service
RAN	Radio Access Network
RB	Resource Block
RT	Real Time
SCS	sub-carrier spacing
SDN	Software-Defined Network
SLA	Service Level Agreements
SMF	Session Management Function
SNMP	Simple Network Management Protocol
SNPN	Standalone Non-Public Network
SPS	Semi-Persistent Scheduling
TAS	Time-Aware Shaper
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TSC	Time Sensitive Communications
TSCAI	Time Sensitive Communication Assistance Information
TSN	Time-Sensitive Networking
TT	TSN translator
UDP	User Datagram Protocol
UE	User Equipment
UL	Uplink
UL CL	Uplink Classifier
UP	User Plane
UPF	User Plane Function
URLLC	Ultra-reliable Low-Latency Communications
VLAN	Virtual LAN
VN	Virtual Network
VPN	Virtual Private Network
WP	Work Package

Executive Summary

This deliverable deals with enhancements for 5G Non-Public Network communications. It mainly focuses on analytical considerations to improve the existing logical 3GPP architecture and its specifications for 5G Non-Public Network deployment scenarios.

First, the advantages and drawbacks of current Standalone Non-Public Network (SNPN) and Public Network Integrated Non-Public Network (PNI-NPN) deployment options defined by 5G Alliance for Connected Industries and Automation (5G-ACIA) are analysed. Based on this analysis, a new deployment option is developed to overcome drawbacks of existing PNI-NPN configurations and to meet OT requirements formulated in WP2, namely, the localization of sensitive OT traffic from industrial applications on-premise without traversing a third party network.

Second, existing industrial Ethernet protocols that are used in wired infrastructure of OT premises are analysed from the viewpoint of their integration issues with 5G. In particular, technical features and challenges of the three protocols that have the most industrial market share, namely, EtherCAT, Ethernet/IP, and PROFINET are studied and ways how they should be integrated with 5G are proposed.

Third, TSN-based aspects for deterministic 5G communications in industrial automation are considered. In particular, a novel Configured Grant enhanced scheduling scheme for 3GPP 5G-TSN integration model with a virtual bridge is described in detail and its performance is validated by means of simulation results.

Forth, other solutions for deterministic 5G communications based on 3GPP standards are analysed to find gaps and resolve some existing 3GPP challenges for industrial automation. In particular, some URLLC enhancements for advanced industrial deployments, integration support of 5G Virtual Network and legacy LAN, and exposure of 5G capability for factory areas are proposed.

Fifth, for OPC UA and 5G interworking new Quality on Demand interface with defined a set of QoS profile levels is proposed. In particular, a procedure to enforce a given level of QoS for a specific OPC-UA connection in the context of an industrial 5G network is presented.

Finally, all observations, gaps, novelties, innovations achieved in this Deliverable in comparison with current State-of-Art are summarised in the Conclusion section.

1 Introduction

Manufacturers are increasingly embracing complex industrial automation digital platforms characterized by vendor lock-in. These platforms integrate sensor-equipped machines that transmit real-time data to cloud analytics engines. The convergence of Information and Communications Technology (ICT) and Operational Technology (OT) empowers manufacturers to address the issue of vendor lock-in, enhance time-to-market, flexibility, security, safety, and sustainability goals. 5G technology plays a vital role in enabling this transformation.

With its ultra-low latency, high reliability, and specialized features like clock synchronization, accurate positioning, and time-sensitive networking, 5G can either replace or supplement traditional industrial wireless and wired infrastructures. It provides a uniform, flexible, scalable, and configurable wireless mobile networks that can be deployed on or off-premises through various options, including leasing from mobile operators, while ensuring security and data privacy.

Industrial automation systems greatly benefit from wireless connectivity facilitated by technologies such as 5G, enabling the widespread adoption of additive manufacturing (AM) across industries. AM allows for the production of customized products, reducing lead times, expediting time-to-market for new designs, and enabling swift responses to customer demands. 5G, unlike its predecessors, is designed with flexibility and versatility, specifically catering to the unique requirements of vertical industries like manufacturing.

5G introduces three service categories: Enhanced Mobile Broadband (eMBB) for high-bandwidth applications, Ultra-reliable Low-Latency Communications (URLLC) for latency-sensitive and highly reliable communications, and Massive Machine-Type Communications (mMTC) for applications involving many devices. The introduced capabilities are continuously improving to make the communication system more suitable for the vertical domain, especially the manufacturing sector. In the domain of industrial automation systems, the 3rd Generation Partnership Project (3GPP) has made significant progress by introducing Time-Sensitive Communications (TSC), Non-Public Networks (NPNs), and 5G-LAN type services. TSC ensures deterministic and isochronous communications with low latency and high reliability. NPNs enable dedicated and independent networks deployed on enterprise premises, offering optimized coverage, low packet latency, enhanced security, and privacy. 5G-LAN type services facilitate LAN communications and Virtual Private Networking (VPN) through 5G systems. 5G and future generations will serve as the central nervous system of future factories, enhancing manufacturing with greater flexibility, mobility, and productivity.

Portable, mobile operating panels empowered by 5G enable on-site workers to connect with various systems and machines, improving convenience and safety. The combination of augmented reality and 5G allows workers wearing data glasses to access real-time information, optimizing machinery monitoring and maintenance. 5G also simplifies the integration of driverless transport systems into manufacturing operations.

The manufacturing industry is currently undergoing a transformative journey fuelled by advanced technologies and the adoption of 5G and beyond. These advancements promise increased efficiency, cost reduction, improved productivity, enhanced customer experiences, and a greener and digitally advanced manufacturing landscape.

1.1 Purpose of the document

This document aims to conduct an in-depth exploration, proposal, and development of innovative solutions for improved 5G and beyond communication and networking. It involves investigating different deployment strategies and introducing technological enhancements for an advanced Cyber-Physical System of Systems (CPSoS) empowered by 5G.

The initial phase of the task involves analyzing standalone and hybrid deployment strategies to determine the optimal approach for deploying 5G NPN. The goal is to strike a balance between reducing network overhead and meeting practical market needs with reasonable costs, taking into account the trade-offs faced by industry customers.

It also involves research activities focused on optimizing industrial mobile communication using advanced 5G solutions. This includes standalone 5G infrastructure, CUPS (Control and User Plane Separation) Architecture, Network Slicing, and leveraging existing multi-access edge computing capabilities to isolate traffic and offload tasks to edge servers. The emphasis, however, lies in integrating IEEE Time-Sensitive Networking (TSN) features to transform 5G into TSN-native end-to-end systems. It enables a wide range of use cases at the factory shop floor, e.g. synchronous communications between sensors, PLCs and actuators used in various industrial automation units or communication between mobile robots, automated guided vehicles (AGVs), and guidance control including process data, video/image, and emergency stop.

To achieve this, additional components might be introduced to integrate fixed Industrial Internet of Things (IIoT) with 5G systems. This includes a scalable IEEE TSN bridge that provides a Network-side TSN Translator (NW-TT) and Device-side TSN Translator (DS-TT). We place particular emphasis on deterministic networking and reliability as essential aspects of industrial communication.

Another area of focus is the research of open and programmable Software-Defined Networking (SDN)-based solutions for monitoring, managing, and proactively adapting the Quality of Service (QoS) and operation of local industrial networks. This involves considering data and service demands, as well as the status of local industrial controllers, to improve scalability, adaptability, and reliability with programmable and data-driven control planes. Interfaces and interaction processes will be defined to ensure effective coordination between open SDN-based local industrial network controllers and the 5G zero-touch management process.

In the future, the task will work on solutions to enhance energy efficiency. This involves intelligent energy-saving techniques and resource allocation in the time domain to optimize energy consumption while meeting industrial reliability requirements. Given the increasing importance of greener industries, improving energy efficiency plays a critical role.

Additionally, it will investigate 5G NPN business models, Service Level Agreements (SLAs), and uptake potentials to gain insights into the economic viability and potential adoption of these solutions.

1.2 Connection with other work packages

This task is associated with several work packages, namely WP2, WP3, WP6, and WP7, each with specific objectives and contributions.

This document receives input from the overall zero-swarm architecture and trial requirements developed in WP2. The focus of this document, as outlined in the table of content, lies in the networking and connectivity layer, with particular attention given to cellular communication systems, including 5G and future generations.

The outcomes of D3.1 are relevant to other tasks within WP3. This document provides general information on deployment models for cellular communication systems, addressing both brownfield (existing infrastructure) and greenfield (new infrastructure) deployments. Practical considerations based on the experience gained in 3GPP and 5G-ACIA are also included.

The activities of WP6 will be supported by T3.1, as it concentrates on the deployment of cellular communication in various trials within the project. This task will provide insights and expertise in this area to contribute to the objectives of WP6.

Furthermore, the outcomes of D3.1 will be utilized to enhance project dissemination, communication, and standardization efforts in WP7. Given the nature of the document, which focuses on cellular communication systems, special attention will be given to relevant 5G pre-standardization activities in related venues, such as 5G-ACIA, ensuring alignment with ongoing industry developments and standards.

1.3 Place in the Zero-SWARM reference architecture

Figure 1 illustrates the topics covered in D3.1 within the reference Zero-SWARM architecture. Several research topics have been identified:

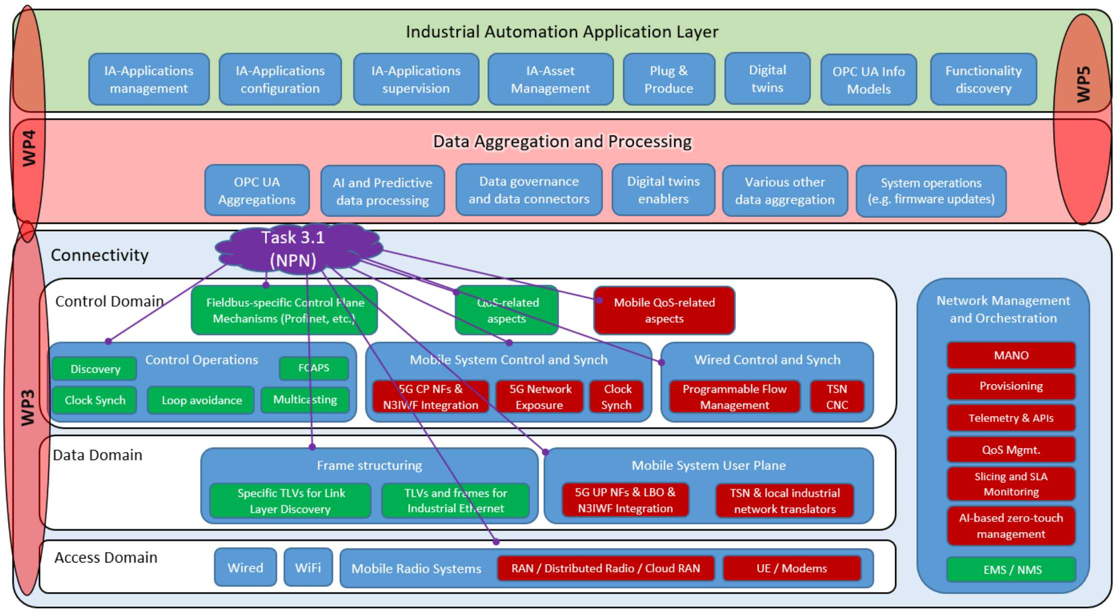

Figure 1 Topics covered in D3.1 within the reference Zero-SWARM architecture

- Greenfield and brownfield scenarios: Greenfield deployments involve new infrastructure, enabling the rollout of state-of-the-art technologies like 5G and beyond. However, data privacy concerns must be addressed, requiring innovative technical solutions and new distribution of responsibilities. Brownfield deployments coexist with existing equipment and systems, often requiring compatibility between different protocols. Wireless communication protocols can act as "wrappers" to enable interoperability.
- Standalone vs. integrated deployment models: Standalone NPNs offer complete control and flexibility but require significant multivendor technology integration skills. Shared RAN deployments provide connectivity and improved data privacy conditions, but sharing the RAN may limit performance and adaptability. Hosted deployments involve hosting the NPN inside a public network, requiring virtualization of network functions and network slicing. The challenge lies in the business viability for MNOs and the need for extensive training for IT personnel.
- Support of cloud-edge continuum: Ensuring interaction between separated edge sites located in different domains (OT production, IT enterprise, and public domains) is a practical challenge. Adaptive production requires processing raw data at the edge of the OT production domain before sending it to analytics applications in the IT enterprise domain. Integration between OT production and IT enterprise domains and a 5G-enabled mesh of User Plane Functions (UPF) across the edge/cloud infrastructure is crucial. The 5G Local Breakout (LBO) capability enables wired or wireless devices/entities in an OT shop floor to access various applications and resources from different edge/cloud systems.

2 Standalone and hybrid 5G NPN deployment strategies

2.1 Existing NPN deployments for different industrial use cases

Different NPN deployment scenarios were defined by 5G-ACIA in [[1]] based on 3GPP specifications and illustrated in Figure 2.

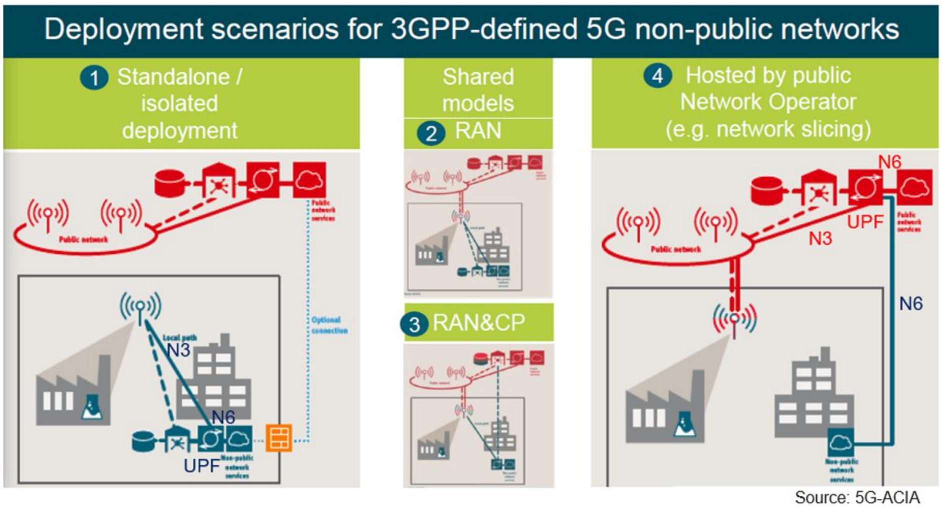

Figure 2 Existing NPN deployment scenarios Error! Reference source not found.

The UPF has a key role in a 5G network to access edge computing system and applications in a data network. In particular, the local UPF takes care of steering the user plane traffic via the N6 reference point towards the data network where the targeted edge applications are located. The local UPF may be a part of the edge implementation as mentioned in [[2]].

Thus, deployment of the UPF and locations of the edge are correlated. 5G-ACIA has defined for integration of industrial infrastructure with 5G networks three different deployment options where the UPF can be physically implemented, namely, in the public network domain, OT production domains, and IT enterprise domain (see Section 4.3 in [[3]]).

The following figure gives an outline of the three deployment options for the physical implementation of the UPF and associated edge computing system based on our observations from **Error! Reference source not found.-Error! Reference source not found.**.

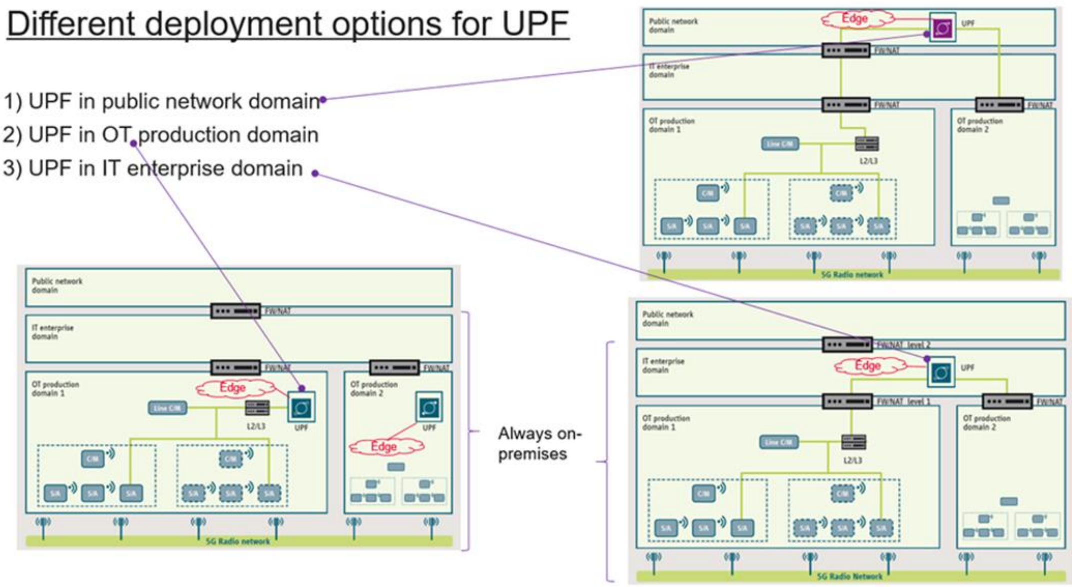

Figure 3 Different UPF and associated edge deployment options for industrial communications

Below we briefly describe the three UPF deployment options summarising their advantages and drawbacks based on [**Error! Reference source not found.** and [[4]].

1) UPF in the public network domain. In this deployment option the UPF is located within the Mobile Network Operator (MNO) network and can be implemented outside or inside of the enterprise premises. That is, all network traffic is routed via the public network to or from the NPN subscribers that are therefore also public network subscribers. This deployment option corresponds to the public network integrated non-private network (PNI-NPN) deployment scenario, namely to the scenario 4 where NPN is hosted by the public network [[1]].

The advantage of this option is that it does not require the integration of the UPF into an enterprise IT/OT network and as a result, it has moderate cost and rapid implementation. Since this option reflects the conventional network layout offered by the MNO it can support typical public network features, e.g., to be suitable for geographically distributed locations.

The drawback of this option is that any traffic from or to devices located in the OT production domain must traverse two layers of gateways (see Figure 3) with firewall (FW) and network address translation (NAT) functionality. It increases complexity related to gateway configurations. Besides, the UPF deployment far away outside of the enterprise premises, e.g., deeply in core network hosted by the MNO, may impact on latency and availability values that are only appropriate for use cases that do not require real-time communication or ultra-high reliability. Also, since the UPF, and correspondingly, the edge infrastructure is located in the public environment it can be questionable from the viewpoint of privacy/security whether it is needed to store the private sensitive information there.

2) **UPF in OT production domain.** In this deployment option the UPF is located within OT production domain and it is always within the enterprise premises.

This option can correspond to standalone non-public network (SNPN), i.e., the scenario 1 in [[1]]. Alternatively, it can be realised as the PNI-NPN scenarios 2 and 3 [[1]], namely, with shared RAN or with shared RAN and control plane.

The advantage of this option is that it provides the best latency values and highest availability since all user plane traffic is localized within the OT production domain and does not traverse FW/NAT gateways.

Since the UPF, and correspondingly, the edge infrastructure is located in closed enterprise environment (i.e., within a private network) those data that are sensitive from the viewpoint of privacy/security can be stored locally and, therefore, safety on-premises.

The drawback of this option is that it requires tight integration of 5G network elements into the existing OT infrastructure that increases complexity of implementation. If an enterprise has multiple segregated OT production domains and a UPF node is in each domain then seamless mobility between domains may be an issue.

3) **The UPF in IT enterprise domain.** This deployment option is similar to the previous one, but since the UPF is implemented on IT layer within the enterprise premises, user plane traffic should traverse one FW/NAT gateway. This option corresponds the same deployment scenarios [[1]] as the option 2, i.e., SNPN or the PNI-NPN with shared RAN and with shared RAN and control plane.

The advantage of this option is that it provides good latency and availability values since all user plane traffic is localized within the enterprise premises and as a result short latency times and high availability can be achieved when the FW/NAT gateway is not a bottleneck. If an enterprise has multiple segregated OT production domains, a UPF node located at the IT layer can provide seamless mobility between domains as a common anchor point that is kept when domains are changed. The same as in the previous option, the sensitive data from the viewpoint privacy/security can be stored locally and safely within the enterprise.

The drawback of this option is that it requires integration of 5G network elements into the existing IT/OT infrastructure that increases complexity of implementation. Due to the FW/NAT gateway (see Figure 3) towards end-devices, latency values are higher than in the previous deployment option that can impact on use cases that require very low latency.

Note that since the UPF and the edge infrastructure locations are mutually bounded (see Figure 3) we assume that the UPF deployment options described above with their pros and cons are also valid for potential edge locations.

2.2 New NPN deployment option for industrial use cases

In this section we propose a new NPN deployment option for industrial use cases that has not been described in 5G-ACIA [[1]]. The purpose of this new NPN deployment option is to address a need from the OT side that was formulated in WP2.

In accordance with the need defined in D2.1, sensitive OT data traffic should be treated locally in OT premises. In particular, it should not be transported through a 3rd party network.

As it is seen from the analysis made in the previous subsection the OT traffic is transported to the 3rd party (i.e. the MNO) when Non-Public Network (NPN) is hosted by the public network operator, i.e. It corresponds to deployment scenario #4 in Figure 2. In this scenario it is assumed that UPF in the public network domain, i.e. in the MNO premises. The advantages and drawbacks of this UPF location are described below Figure 3, the case 1.

The main advantage is that it does not require the integration of the UPF into an enterprise IT/OT network and as a result, it has low cost and rapid implementation for OTs. It can be attractive for small and medium enterprises (in contrast to a standalone isolated NPN, option 1 in Figure 2, that is more for large ones).

One of main drawbacks of this deployment scenario is that since the UPF is located in the public environment it can be questionable from the viewpoint of privacy and security.

The problem is even though OT "sensitive" data from OT own applications can be stored and processed locally, e.g., in on-premise local edge, they are anyway **transported via third party** (e.g. MNO) infrastructure as illustrated in Figure 4:

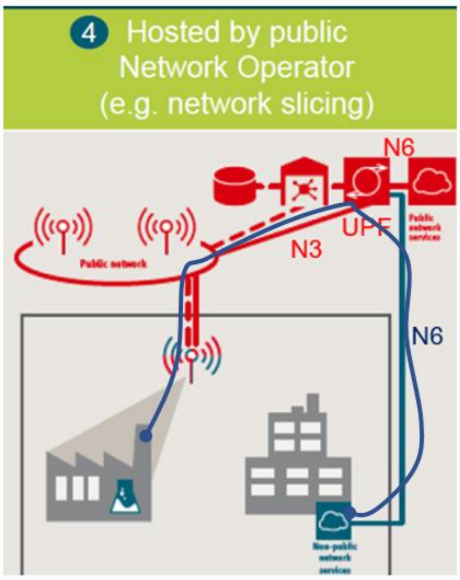

Figure 4 Error! Not a valid bookmark self-reference.. **OT data traffic path in the NPN deployment hosted by MNO (all OT traffic goes via MNO including sensitive data)**

Since private sensitive information needs to be kept locally, but it is transported outside of OT premises the option#4 (Figure 2) is not suitable to address this OT need.

Project funded by Horizon Europe, Grant Agreement #101057083

In this section we propose new deployment option that can meet this requirement and localize sensitive OT data traffic on-premise.

The solution is based on the introduction of an Offloading Point (OP) to breakout sensitive OT data and localize it on-premise.

The OP functionality is to offload sensitive OT data (e.g. OT edge-related traffic) localizing them on-premise, but the rest of traffic (not sensitive) should go via the UPF of MNO to external networks or services if required (the same path as in Figure 4).

The OP is located on the path between a base station (gNB) and the UPF, but within the enterprise, e.g., it can be located in the gateway of the enterprise to the external environment (i.e. public network domain) as illustrated in Figure 5 below.

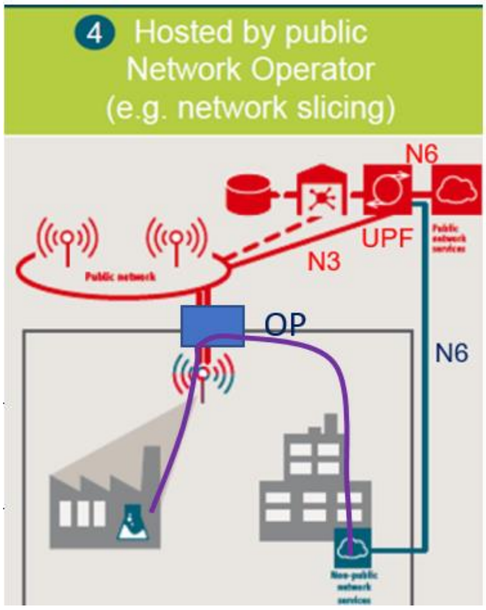

Figure 5. Sensitive OT traffic data path localized in OT premises by means of offloading point

Note that the Offloading Point is a non-3GPP entity, but more an “ETSI entity” that has traffic filters and makes traffic inspection. The idea is based on considerations in [[5]]. That is, the specific feature of the OP is that the rules for traffic filter and inspection for the OP are received from the edge orchestrator via the ETSI’s Mp2 interface (in contrast to 3GPP where the UPF receives the policy and rules for data path forming from the Policy Control Function (PCF)/Session Management Function (SMF) as defined in 3GPP TS 23.501).

If the traffic destination is OT wireless devices, OT wired devices, OT edge applications, etc. then the traffic is considered as sensitive and steered by the Offloading Point by means of received rules and filters from the edge orchestrator to be processed and kept on-premise.

The steering should be done from the N3-interface of MNO. The Offloading Point can be an SDN-based switch using P4 including the GTP User Plane (GTP-U) encapsulation and decapsulation functions and the forwarding and steering functions as mentioned in [[6]].

Note that that existing 3GPP uplink clarifier (UL CL) feature in 3GPP TS 23.501 does not address this problem. This feature from the viewpoint of its applicability for industrial communications is illustrated in Figure 6:

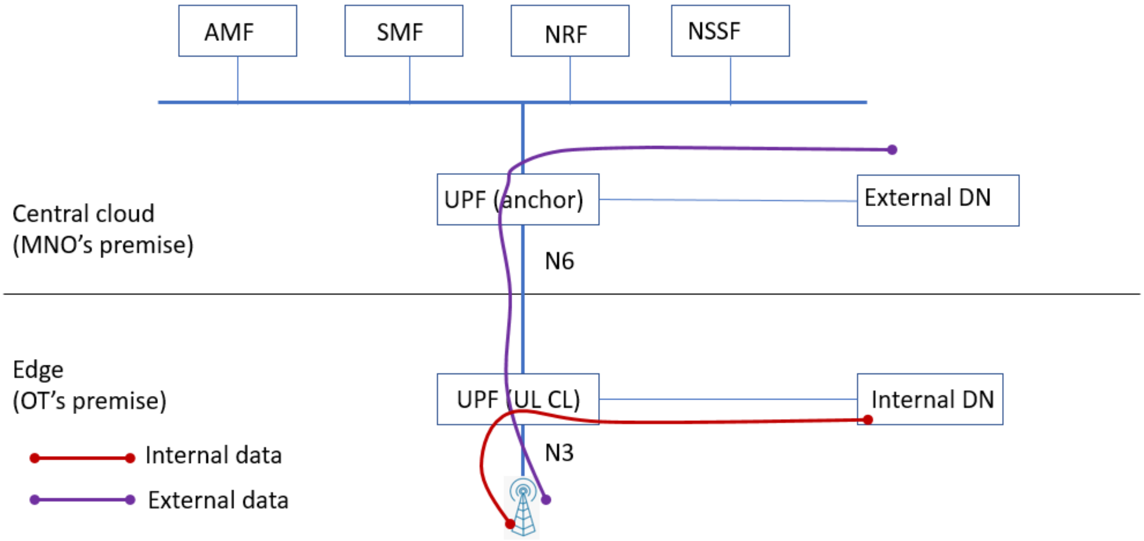

Figure 6 Data traffic offloading to local data network by means of 3GPP uplink classifier

In accordance with 5G-ACIA the deployment option #4 (Figure 2) should not have the UPF on-premise, that is, only edge can be on-premise. It is not a case for the 3GPP UL CL feature shown in Figure 6 that requires from OT to deploy the UPF locally under control of the MNO (note also that the UPF is very expensive network equipment).

Besides, the OT need from WP2 states that the **private sensitive information** should be **processed, controlled and stored on-premise**, i.e. it should not have any 3rd party involvement (e.g. MNO).

That is, even though in a case if the UL CL UPF is installed on-promise, anyway the traffic will be controlled by the third party (i.e. MNO) that is not appropriate for OT from the viewpoint of privacy and security. For instance, traffic filters and templates for traffic inspection are sent by the PCF, i.e. MNO is aware about these filters and templates and can judge which traffic is offloaded.

In case of this new deployment option, we propose that the Offloading Point uses traffic filters to make traffic inspection that are received via the Mp2 interface [[2]] from the edge orchestrator that belongs to the OT and is located in the OT premises.

2.3 Summary

As a summary of this section, a new NPN deployment option has been proposed in addition to the existing ones defined in the 5G-ACIA white paper **Error! Reference source not found.**[1]]. In comparison with the 5G-ACIA deployment option #4 (see Figure 2) the new NPN deployment option can support sensitive traffic "localisation", in particular, sensitive OT data is not only stored locally, but also not transported via a third party (the UPF of MNO) that is not a case in the existing 5G-ACIA option #4 **Error! Reference source not found.**]. Also, the proposed deployment option can support industrial **use cases with delay critical** requirements (e.g. closed loop control and mobile robots) since data traffic transmission path is shorter since it is on-premise localised (please check data paths in Figures 4 and 5 for comparison). Moreover, the proposed deployment option has minimal cost in comparison with all 5G-ACIA deployment options since it does not require the expensive MNO's UPF equipment installation in as, e.g. in the options 1-3 in Figure 2 (only low-cost SDN-based P4 programmable switch for traffic offloading is needed).

3 Integration of industrial Ethernet protocols with 5G networks

3.1 EtherCAT, Ethernet/IP, PROFINET and their integration aspects with 5G

A number of Industrial Ethernet protocols have been introduced since the early 2000s to meet the communication requirements on the factory floor, thanks to the ability to achieve integration of control and management through the same wired infrastructure, and much faster speed of Ethernet in comparison with traditional fieldbus. Figure 7 shows the market penetration of the most popular protocols employed in automation, where the industrial Ethernet protocols rank first with 66% share of the whole market and a 10% annual growth. In the following, we present three of such protocols and discuss the challenges of the integration with 5G systems.

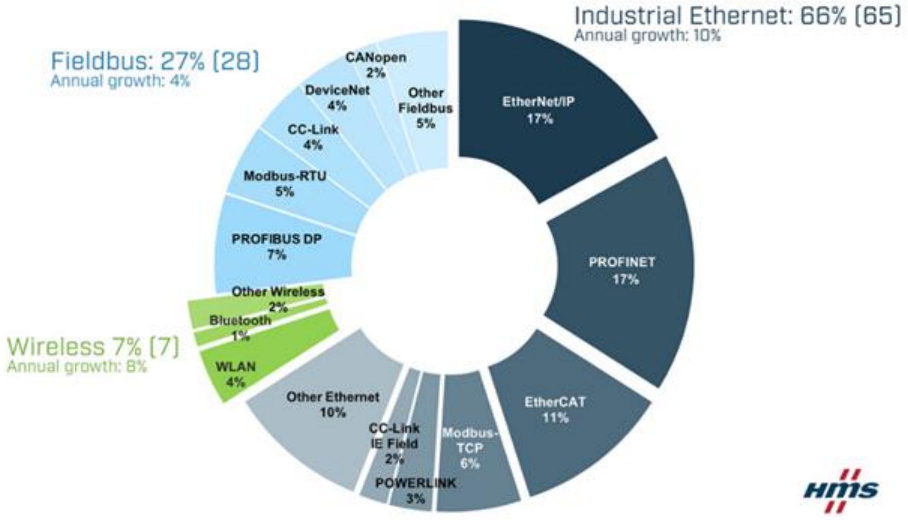

Figure 7 2022 Industrial Network Market Shares (source: HMS)

The three industrial protocols that have the most market share, namely, EtherCAT, Ethernet/IP, PROFINET are considered in this section in the context of their integration aspects with 5G.

3.1.1 EtherNet/IP

EtherNet/IP (IP: Industrial Protocol) is an application-layer industrial Ethernet protocol developed and maintained by Open Device Vendors Association (ODVA) and promoted primarily by Rockwell Automation. It is member of a family of networks implementing the Common Industrial Protocol (CIP™), a technology for organizing and sharing data in industrial devices. Specifically, EtherNet/IP specifies how CIP communication packets can be transported over standard Ethernet and TCP/IP technology. It is classified as Class 1 real-time Ethernet protocol according to IEC 61 784-2 as the real-time mechanisms are embedded in the top layer. Figure 8 shows the ISO/OSI model of EtherNet/IP and two more protocols supporting CIP, namely, ControlNet and DeviceNet. As illustrated in the figure, EtherNet/IP consists of the Ethernet physical layer and the Carrier Sensing Multiple Access and Collision Detection (CSMA/CD) at the Data Link layer, while the TCP/IP suite is employed for the network and transport layer in order to encapsulate CIP messages. At the higher layers, a number of CIP-specific services, libraries and interfaces are used to handle the communication of different devices.

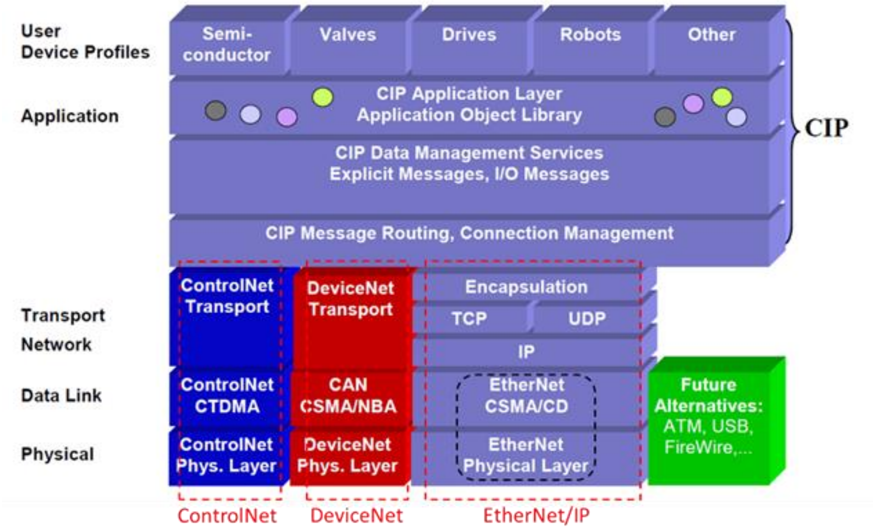

Figure 8: ISO/OSI model of EtherNet/IP

Two types of CIP message types are available:

- **Implicit or I/O**, suitable for cyclic real-time producer/consumer communication. Implicit messages are carried over UDP allowing the delivery of packets from a single producer to many consumer (multicast transmission). An example of application relying on implicit messaging is real-time control, where fast transmission speed and low latency are essential.
- **Explicit**, suitable for acyclic non-real time client/server communication. Explicit messages are transmitted by TCP and include destination, source and connection information. As a result, explicit messaging is less efficient than implicit messaging, yet it offers a high degree of flexibility. Example of applications relying on explicit messaging are diagnostics and data configuration.

EtherNet/IP devices are classified into three classes according to the network communication capabilities:

- **Scanner class** (e.g., PLCs and controllers), which maps network input and output variables in cycles, defined through the update time.
- **Messaging class** (e.g., diagnostics, network configuration tools, SCADA and HMI systems), which supports explicit messaging rather than real-time I/O data.
- **Adapter class** (e.g., sensors, valves, gateways), which offers device-specific features.

Figure 9 shows how these three classes interact with each other in an EtherNet/IP network deployment.

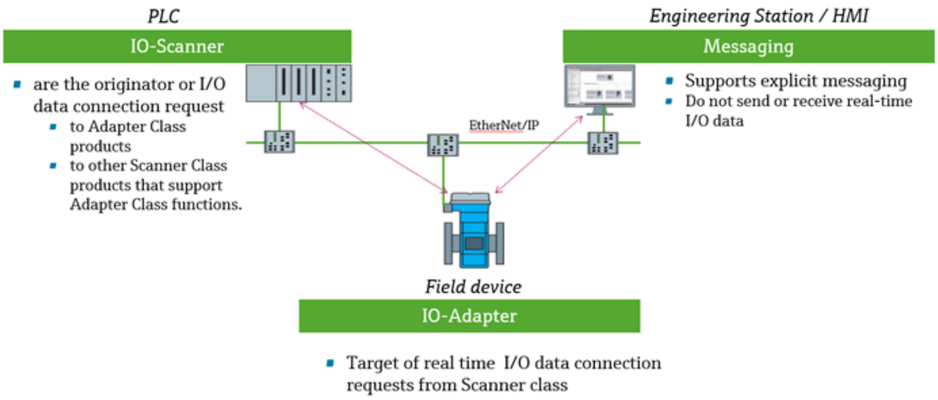

Figure 9 Interactions between EtherNet/IP device classes

Project funded by Horizon Europe, Grant Agreement #101057083

The general QoS approach to EtherNet/IP calls for devices to mark their packets with a priority value. Starting from edition 1.6, EtherNet/IP specifications defines the behavior of the networks in relation to the QoS. Two important QoS mechanisms are considered, i.e., Differentiated Services (DiffServ) and IEEE 802.1D/Q. In general, switches and routers must be able to differentiate real-time traffic from non-critical packet streams, i.e., implicit vs. explicit messages.

DiffServ is defined in IETF RFC 2475 and is a model for providing different treatments of traffic classes based on their characteristics, tolerance to delay, jitter and packet loss, operating at layer 3. Networking nodes process packets based on class of traffic defined by DiffServ Codepoint (DSCP) and associated Per-Hop Behavior (PHB) characteristics. The DSCP values used by EtherNet/IP are the ones defined as "local use" in the IETF RFC 2474 and appear in the ToS field of the IP header.

IEEE 802.1D/Q defines an Ethernet frame format allowing inclusion of VLAN ID and priority. The 802.1Q frame has EtherType of 0x8100 and a 4-byte prefix between the Source and the Type fields. The tagged frame defines a 3-bit field to specify 8 priority levels, further specified in 802.1D, also called CoS (class of service), where priority 7 is the highest and 0 the lowest.

Table 1 provides the mapping between traffic type and the DSCP or 802.1D priority. Furthermore, the EtherNet/IP default behavior is to mark packets with DSCP values, yet devices may optionally support sending and receiving 802.1Q frames with the corresponding priority values.

Table 1 Mapping between traffic type and DSCP or 802.1D priority.

Traffic Type	CIP Priority	DSCP	802.1D Priority[1]	CIP Traffic Usage (recommended)
PTP Event (IEEE 1588)	n/a	59 ('111011')	7	n/a
PTP General (IEEE 1588)	n/a	47 ('101111')	5	n/a
CIP class 0 / 1	Urgent (3)	55 ('110111')	6	CIP Motion
	Scheduled (2)	47 ('101111')	5	Safety I/O I/O
	High (1)	43 ('101011')	5	I/O
	Low (0)	31 ('011111')	3	No recommendation at present
CIP UCMM CIP class 3 All other EtherNet/IP encapsulation messages	All	27 ('011011')	3	CIP messaging

[1] Sending 802.1Q tagged frames is disabled by default

Being a protocol compliant with Ethernet and TCP/IP, EtherNet/IP can be easily supported by a 5G system. Specifically, the IP packet carrying a CIP packet is encapsulated in a 5G PDU (packet data unit) and delivered to the destination by going through the 5G network, down to the 5G UE connected to target EtherNet/IP device. Figure 10 shows the EtherNet/IP stack and the 5G stack as well as the EtherNet/IP packet embedded into a 5G PDU. Finally, a suitable mapping between the EtherNet/IP QoS priority and the 5QI (5G QoS Identifier) must be implemented in order to meet the requirements of each packet in terms of delay budget and error rate.

Project funded by Horizon Europe, Grant Agreement #101057083

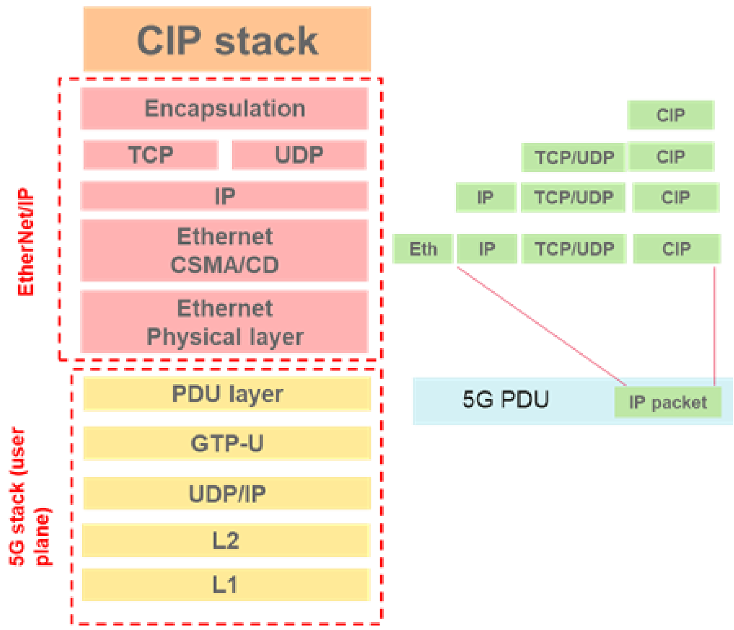

Figure 10. Integration of EtherNet/IP with 5G.

3.1.2 EtherCAT

Ethernet for Control Automation Technology (EtherCAT) was first developed in the early 2000's by Beckhoff and it is also called EtherCAT Device Protocol to differentiate it from EtherCAT Automation Protocol. It is governed by a number of standards, i.e.:

- IEC 61158 – defining how EtherCAT adapts the OSI communication model
- EtherCAT Fieldbus Network
- IEC 61784-2 – defining profiles for specific device classes
- IEEE 802.3 Ethernet working group

EtherCAT is a high-performance Ethernet-based fieldbus network allowing to transfer information at high speed and deterministically. The protocol employs the master/slave principle to control access to the medium. The master node (typically the control system) transmits Ethernet frames to the slave nodes (e.g., sensors and actuators), which can extract data from and insert data into these frames. The slave nodes have two EtherCAT ports: one port is assigned as input (to receive telegrams) and the other one as output (to forward telegrams). Also, slaves are equipped with a specific hardware called EtherCAT Slave Controller (ESC) which is responsible for processing and forwarding the frames to the next slave at high speed.

As shown in Figure 11, the process data of the whole network is carried together in one or more Ethernet frames following the "summation frame principle", as opposed to the individual frame approach where every frame carries process data for only one device. Furthermore, the standard Ethernet packet is not received, interpreted and copied at every node of the network. Instead, slave devices process frames "on the fly", reading and inserting data while the frame is passing through the device. For time synchronization, EtherCAT relies on a clock synchronization mechanism known as distributed clock (DC). All the DC-enabled slaves in the network are synchronized with a common timing reference under direct control of the master.

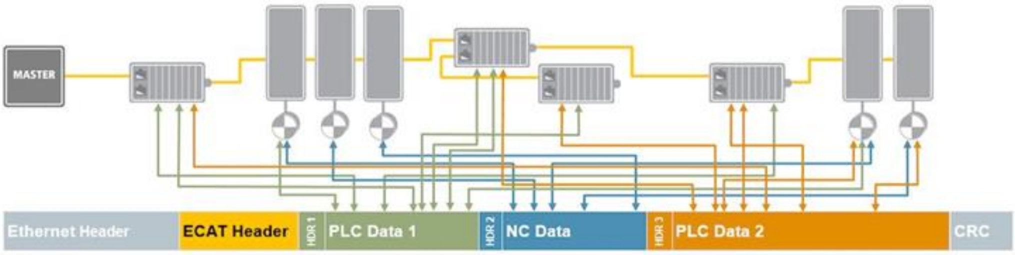

Figure 11. Illustration of the summation frame principle.

Figure 12 shows the ISO/OSI model of EtherCAT. The physical layer is based on the IEEE 802.3 standard for 100Base-TX, i.e., the Ethernet physical layer. On the other hand, data link and application layer are specifically designed for motion control applications, while TCP/IP is optional. EtherCAT frames can be transported in two ways:

- Directly over an Ethernet frame with EtherType 0x88A4, in cases where maximum performance is required and the EtherCAT nodes are all in the same Ethernet subnet. Up to 65,535 nodes can be addressed in the same EtherCAT segment.
- Over a UDP payload through UDP port 0x88A4, in cases where IP routing is required to interconnect EtherCAT nodes belonging to different subnets and packet latency can be tolerated.

No.	OSI-Layer	EtherCAT	
7	Application Layer	http*, ftp*	Cyclic Data Exchange Mailbox Acyclic Data Access
6	Presentation Layer	-	-
5	Session Layer	-	-
4	Transport Layer	TCP*	-
3	Network Layer	IP*	-
2	Data Link Layer	Mailbox/Buffer Handling, Process Data Mapping, Extreme Fast Auto-Forwarder	
		Ethernet MAC	
1	Physical Layer	100BASE-TX, 100BASE-FX	

*optional

Figure 12. ISO/OSI model of EtherCAT.

Communication is always initiated by the EtherCAT master (e.g., a PLC) sending a telegram which passes through all the EtherCAT slaves. The last slave of the communication chain can detect that no other node is attached to the same chain, and can return the telegram to the master. Each slave receiving the telegram takes the portion of data intended to it and writes its data back into the Ethernet frame. This process takes place on the hardware level, thus ensuring minimum processing and forwarding delay. On the return path, the telegram is directly sent to the master without being processed in the slaves.

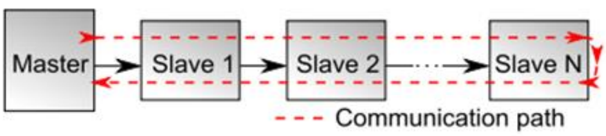

Figure 13: EtherCAT line topology.

Project funded by Horizon Europe, Grant Agreement #101057083

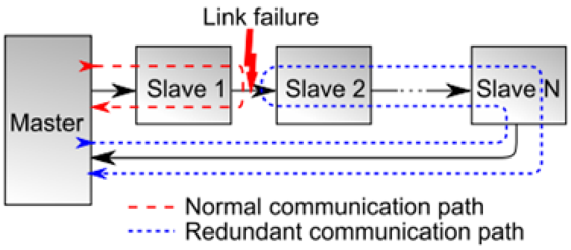

Figure 14: EtherCAT ring topology with redundancy.

EtherCAT flexibility allows to employ and combine multiple topology variants, including line, star, tree and ring (or daisy chain). Figure 13 and Figure 14 shows the EtherCAT line topology and ring topology, respectively. The ring topology provides a failover mechanism for addressing link failures between slaves. The same telegram is transmitted over two ports of the same master and the secondary port is connected to the last slave of the chain. This way, the master can ensure that the telegram will reach all the slaves.

Due to the fact that EtherCAT adopts a non-switched network architecture where no Ethernet switches are employed, telegrams do not require any special QoS treatment. Packets are quickly processed and forwarded through the ESC hardware and no buffering or classification is performed. As a consequence, EtherCAT is one of the most efficient industrial Ethernet protocols, ensuring real-time packet delivery.

In order to evaluate how EtherCAT can be integrated with 5G cellular networks, the following considerations have to be made:

- Unlike EtherNet/IP, EtherCAT does not require the TCP/IP stack to transfer telegrams between slaves.
- EtherCAT telegrams do not contain a source and destination address since all the nodes in the same segment receive, process and forward them.
- EtherCAT is not compatible with traditional switched networks, where packets are forwarded by switches and routers. The introduction of intermediate Ethernet switches is not recommended since they can increase the propagation delay and the communication jitter.
- An Ethernet switch linking two EtherCAT devices in the same segment must implement a Link Loss Forwarding (LLF) or Link Loss Carry Forward (LLCF) mechanism, which allows a switch to automatically shut down upstream ports as soon as a signal loss is detected on the downstream ports. Otherwise, the upstream device would not be able to detect a physical layer issue and would not shut down the port connected to the Ethernet switch. This is also pointed out in the ETG.1600 Installation Guidelines.
- Employing a 5G system as a logical bridge in an EtherCAT segment may be more suitable for the case of a star topology. The 5G bridge would then replace an EtherCAT junction, yet it will need to be appropriately instructed in order to sequentially route the telegrams from the master to the last slave and then return them to the master. This is necessary as telegrams do not contain source and destination address.

3.1.3 PROFINET

PROFINET (Process Field Network) is an industrial Ethernet communication standard in the context of factory automation. It is designed to provide real-time communication between industrial devices such as programmable logic controllers (PLCs), distributed control systems (DCS), actuators, and sensors.

PROFINET enables the exchange of data and control information over standard Ethernet networks, allowing for seamless integration of various devices from different manufacturers.

PROFINET utilizes standard Ethernet technology, which allows for high-speed data transmission, compatibility with existing IT infrastructure, and easy integration with other Ethernet systems.

On the other hand, it supports real-time communication, supporting the deterministic and time-critical control and monitoring of industrial processes. This is fundamental in applications where precise synchronization and fast response times are demanded.

Regarding its flexibility, it is a protocol that can be adapted to multiple industrial systems. It supports different network topologies such as line, ring and star configurations. Moreover, it allows the integration of systems with various scales with thousands of devices. However, as it is technologies developed over Ethernet, it has lack of mobility which can impact in the flexibility of the whole factory floor. Especially, in scenarios where mobility is crucial.

The mentioned advantages cause that PROFINET had been adopted in various industries such as manufacturing and automotive. Its versatility, real-time capabilities, and extensive device support make it a robust communication solution for industrial automation and control systems. For these reasons, a study of the integration of PROFINET in 5G networks is a goal in the Zero-SWARM frame.

3.1.3.1 PROFINET Classifications

3.1.3.1.1 Real-Time (RT) PROFINET

Real-time PROFINET has been widely used for applications that require time-critical control, synchronization, high-speed data exchange, deterministic communication, precise timing, and low latency response.

Time Synchronization: PROFINET supports time synchronization mechanisms to achieve precise coordination among devices in a network. It utilizes the Precision Time Protocol (PTP) to synchronize clocks across the network, enabling accurate time stamping of data and ensuring synchronized actions within the system. Time synchronization becomes fundamental in applications where coordinated motion control or event-driven processes are critical.

High-Speed Data Exchange: PROFINET offers high-speed data exchange, which result in a fast and efficient communication between devices, enabling real-time monitoring and control of industrial automation systems.

Deterministic Communication: Real-time PROFINET provides deterministic communication, where data transmission occurs with predictable and consistent timing. It ensures that critical control data is delivered within predefined time constraints, enabling precise synchronization and coordination of industrial processes.

Redundancy and Availability: Real-time PROFINET incorporates redundancy schemes to improve network resilience and fault tolerance. These strategies, such as Media Redundancy Protocol (MRP) or Parallel Redundancy Protocol (PRP), enable seamless switch-over to redundant paths or devices in case of failures. This redundancy attends to guarantee continuous operation and minimizes downtime in critical industrial applications.

3.1.3.1.2 Isochronous Real-Time (IRT) PROFINET:

Isochronous Real-Time (IRT) is a specific real-time extension of PROFINET that provides enhanced determinism and synchronization capabilities. Some of the most significant features are presented as follows:

IRT PROFINET provides strict determinism by guaranteeing a maximum jitter for real-time data transmission. It allows for highly predictable and consistent communication timing, which is critical for applications with stringent timing requirements.

On the other hand, it supports precise time synchronization across the network using the IEEE 802.1AS-Rev standard, which is based on the Precision Time Protocol (PTP). This enables synchronized actions and coordinated motion control in distributed automation systems.

IRT PROFINET introduces isochronous communication channels, which provide dedicated, time-sliced slots for periodic data transmission. These channels are reserved exclusively for time-critical data, such as fast control loops or synchronized motion control, ensuring minimal delays and reliable delivery.

Additionally, it enables the reservation of network bandwidth for isochronous communication. This ensures that isochronous data always has sufficient network resources and is not affected by other non-real-time traffic, enhancing the determinism and reliability of critical data transmission.

These features of Isochronous Real-Time PROFINET make it suitable for applications that demand ultra-fast response times, synchronized motion control, and high precision in time-critical industrial automation processes.

3.1.3.1.3 Non-Real Time (NRT) PROFINET:

Unlike PROFINET mainly focuses on real-time communication for time-critical control and monitoring. However, there is also a non-real-time version that supports additional features beyond the strict real-time requirements. These non-real-time capabilities include: configuration and diagnosis, acyclic communication, and web-based management interfaces, among others.

Its configuration capabilities allow not only to configure the network devices but also the management of I/O modules and to setup of communication parameters. Additionally, the diagnosis functions enable monitoring and troubleshooting of network devices, aiding in the maintenance and optimization of the network.

Additionally, it uses a cycle communication in non-real-time data transmission, configuration requests, and other asynchronous exchanges of information between devices. It enables device-specific data transfers, such as diagnostics, alarms, firmware updates, and network management tasks.

On the other hand, NRT PROFINET supports web-based management interfaces, which provide a user-friendly manner to access and configure network devices. This allows users to conveniently access device parameters, diagnostic information, and configuration settings using a standard web browser.

These non-real-time features expand the capabilities of PROFINET, however they are not suitable for time-critical control applications. The real-time features of PROFINET, with specific time synchronization and deterministic communication, must be considered in applications that require precise timing and rapid response.

As the scope of this deliverable is to explore how well-established Ethernet protocols can be integrated with 5G networks, PROFINET IRT will not be covered. This is mainly due to the fact that PROFINET IRT implements significant changes at layer 2 (MAC level) and requires dedicated hardware for packet processing and forwarding. This deliverable considers Real Time PROFINET as starting point to provide 5G integration guidelines.

3.1.3.2 PROFINET Technical details

This subsection describes the main technical aspects that are fundamental to take into account in the PROFINET and 5G integration. It describes the type of devices in common PROFINET networks, the protocol stack, the PROFINET datagram and the supported topologies.

In general, the devices in PROFINET are classified into three types: supervisor, controller and device. The supervisor is employed for commissioning and diagnostic data collection. Such device can read and write internal diagnostic data of a PROFINET stack or diagnostic data provided by the application program of a device. On the other side, the controller is typically integrated in PLC, motion controller, etc. A controller hosts the configuration data of its dedicated IO-Devices. Also, it initiates the communication to the IO-Device by exchanging and managing cyclic process data. The device is usually integrated into field devices, sensors, motion-drive systems, etc. A device provides cyclic input data and consumes cyclic output data to/from the controller, respectively. Furthermore, it hosts diagnostic data and issues alarms. Figure 15 shows an overview of the three types of devices.

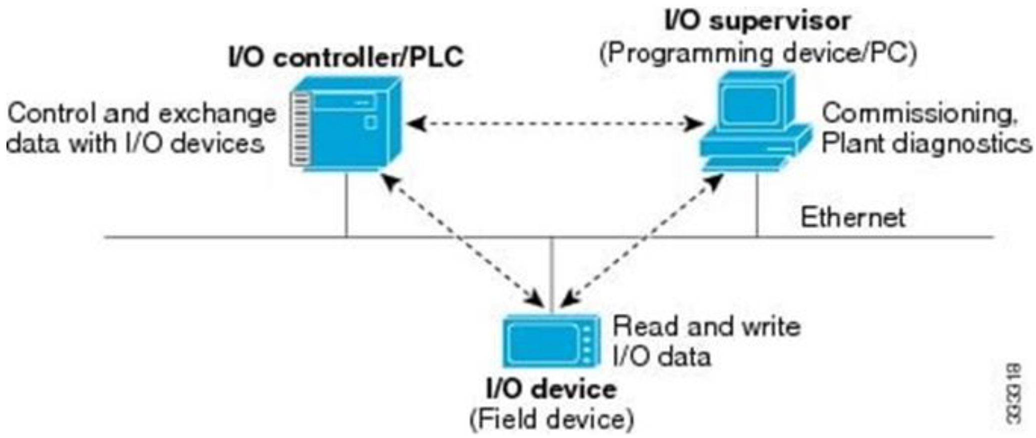

Figure 15. Different types of devices communicating in a PROFINET network

It is crucial to mention that any data exchange is embedded into a logical link called Application Relation (AR). Within an AR, multiple Communication Relations (CR) specify the type of data exchanged: IO data AR (for cyclic data exchange), record data AR (for acyclic data exchange), and alarm CR (for sending alarms).

An IO device can have multiple ARs established from different IO controllers as in the case of shared devices. Furthermore, PROFINET allows to employ and configure multiple IO controllers to access the same data originated from IO devices. Figure 16 illustrates the concept of having multiple CRs between a controller and a device.

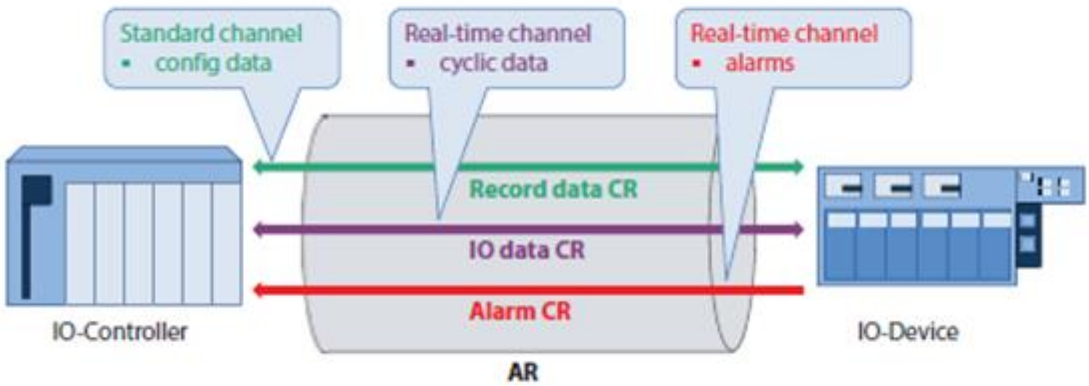

Figure 16. Multiple CRs running between a controller and a device

PROFINET reuses the most popular communication protocols designed for transmitting and receiving information packets. As shown in Figure 17, layer 1 and 2 of the ISO/OSI model corresponds to the unmodified Ethernet protocol. However, PROFINET implements a direct cross-layer channel between Ethernet and the application layer, which avoids packets to be processed by the TCP/IP stack with consequent latency reduction. By skipping the TCP/IP stack, PROFINET can achieve the real-time performance required by most of the automation applications. At the same time, it ensures compatibility with non-real-time applications by incorporating the TCP/IP stack. This also means that any type of traffic travelling on an IP network can be carried over a PROFINET network. As an example, a client running on a standard PC can access a controller through a web browser thanks to the HTTP support.

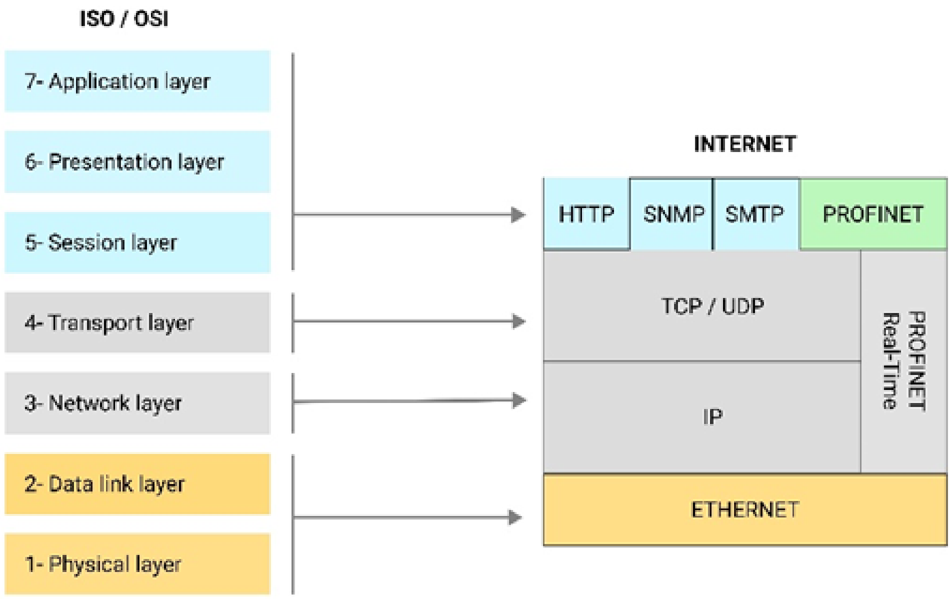

Figure 17. PROFINET protocol stack and mapping with the ISO/OSI model

As shown in Figure 18, a PROFINET datagram has a standardized format according to IEEE 802.3 and can feature two different EtherTypes: “0x0800” for IP, and “0x8892” for PROFINET real-time. Namely, EtherType indicates when the Ethernet frame is a PROFINET real-time Ethernet frame and has to be directed to the PROFINET application without involving the TCP/IP layers.

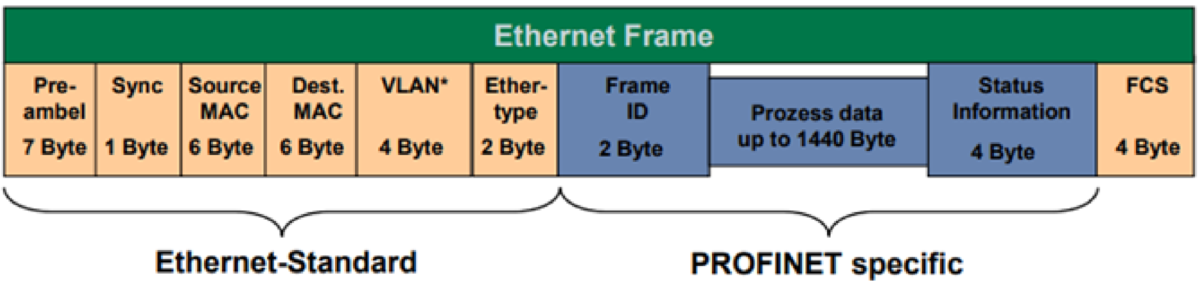

Figure 18. Structure of a PROFINET telegram

Furthermore, PROFINET supports multiple network topologies: line, star, tree, redundant ring, and wireless, which are compared in table 2 and briefly introduced as follows.

Line: This is deployed through PROFINET devices with built-in switch, allowing to connect devices in a line topology without external switches. Such topology is adopted for applications in extensive automation plants such as conveyor belts.

Star: This is deployed through built-in or standalone switches and allows to ensure a certain degree of redundancy with respect to the line topology. The star is suitable for areas with limited geographical extension.

Tree: It is created by connecting multiple star-shaped networks to form one single network. Such topology is a typical example for an automation plant being grouped into different manufacturing islands.

Redundant ring: A ring topology can be implemented by adopting two protocols, i.e., Media Redundancy Protocol (MRP) and Media Redundancy for Planned Duplication (MRPD). The ring is usually deployed to prevent a line topology to get interrupted due to a device outage.

Wireless: PROFINET devices can be wirelessly connected by Wi-Fi or Bluetooth.

Table 2 below presents overview of the different PROFINET topologies.

Table 2. Overview of the PROFINET topologies

Topology	Description	Benefits	Challenges
Line	Most PROFINET devices have at least two ports, with such ports being part of a built-in switch in the device. Built-in switches allow users to connect devices in a line topology without external switches.	Built-in switches in PROFINET devices are usually available, external switches are optional. Less cables needed.	If a node in the line fails, downstream nodes lose communication. User must account for line depth limitations.
Star and tree	Built-in or standalone switches allow for star/tree topologies.	Efficient topology for multiple nodes in close proximity	If a central switch fails, communications to all connected nodes will be affected. Additional cables and external switches
Ring	PROFINET allows you to implement a ring topology by managing it with two defined media redundancy protocols: Media Redundancy Protocol (MRP) and Media Redundancy for Planned Duplication (MRPD)	Media Redundancy: In case of a failed cable or node within the ring, operations may continue.	Devices in the ring must support PROFINET media redundancy. Additional cabling are required. Additional configuration of the devices that make up the ring are needed.
Wireless	Wi-Fi or Bluetooth can become part of a PROFINET network backbone.	No need for cables or additional hardware	Limited speed. Installation requires a depth analysis of the wireless conditions (interference, fading, etc.)

3.1.3.3 5G and PROFINET integration aspects

In order to assess whether and how a PROFINET network can be integrated with 5G systems, the following considerations must take into account multiple requirements such as (but not limited to):

QoS support: Ethernet frames carrying real-time PROFINET datagrams that are usually marked with a VLAN priority 7 (high-priority traffic) to allow fast packet switching. As a result, the 5G system must be able to map such VLAN priority to an appropriate 5QI to ensure an appropriate level of packet QoS.

LLDP support: The Link Layer Discovery Protocol (LLDP) is a layer-2 protocol standardized as IEEE 802.1AB used by PROFINET field devices to detect neighbours in the networks. LLDP messages contain a special EtherType. This allows the respective neighbour to be explicitly identified and the physical structure of the network to be determined. As a result, a 5G system must be able to process received LLDP frames and send them to each 1-hop neighbour. A tentative solution is the implementation of the required procedures to manage LLDP packets in the 5G edge.

L2 multicast: PROFINET employs different protocols requiring multicast packet transmission, i.e., DCP (discovery and configuration protocol), LLDP, MRP and PTCP (Precision Transparent Clock Protocol). This implies that the 5G system must be able to process and forward Ethernet frames addressed to specific multicast groups.

SNMP support: Simple Network Management Protocol (SNMP) is the mandatory network management protocol for PROFINET devices belonging to the conformance class B (CC-B). As a consequence, a 5G system compliant with PROFINET CC-B must be able to support SNMP traffic.

Switch embedding PROFINET IO device functions: PROFINET devices belonging to CC-B are also equipped with a PROFINET stack which allows to send diagnostics data (as alarms) directly to the PROFINET IO controller. Furthermore, such devices have an associated descriptor file called GSD (General Station Descriptor) that contains data for engineering and information exchange between the controller, the supervisor and the devices. In other words, the GSD file describes the properties of the device and contains all this information required for configuration, e.g., device identification number, error text for diagnostics, communication parameters for IO devices (minimum cycle time, reduction ratio, watchdog time, etc.). These requirements imply that a 5G system must also be equipped with a PROFINET stack and must have an associated GSD file to support network diagnostics from a controller through PROFINET-specific mechanisms.

3.1.4 Summary

As a summary of this section, an overview of the three most popular industrial Ethernet protocols has been presented with the goal of assessing the efforts required to integrate them with 5G. Table 3 below provides a brief summary of the main features of each protocol and outlines the potential 5G enhancements required to achieve the integration objective.

Table 3. Main aspects and 5G enhancements required for integration

Protocol	Main features	5G enhancements required for integration
EtherNet/IP	• Flexible and compatible protocol (compatible with TCP/IP and Ethernet) thanks to the use of off-the-shelf components and its object-oriented basis • EtherNet/IP networks lean towards star topologies, but can often be found in ring and tree topologies to enhance redundancy. • EtherNet/IP has limited performance in terms of speed and latency due to usage of the entire TCP/IP stack if compared with EtherCAT and PROFINET.	• The flexibility and TCP/IP compatibility of EtherNet/IP results in a very minimum effort to integrate with 5G. • The 5G system must implement suitable mechanisms to map the EtherNet/IP QoS priority to the specific 5QI.
EtherCAT	• Only employed in non-bridged networks (not suitable for L3 routing). • EtherCAT telegrams linearly traverse a network of slaves and then return to the master which started the communication.	• EtherCAT is not as flexible as EtherNet/IP due to the fact that the TCP/IP support is only optional and not employed for cyclic data exchange • A 5G system can be used to connect a master with the first

	• A specialized hardware called ESC is employed on each slave to enable very fast packet processing and forwarding. • EtherCAT can support multiple network topologies; however, line and ring are the most popular ones.	slave of the EtherCAT segment, but the 5G system must implement an LLF mechanism to address possible link disruption. • Interconnecting all the slaves through 5G is not yet viable due to the very-low latency performance of a traditional wired EtherCAT network which makes use of specialized hardware.

PROFINET	• Support of multiple network topologies.
	• The PROFINET real-time protocol achieves cyclic communication with cycle time lower than 10 ms by forwarding Ethernet frames to the application layers without involving the TCP/IP stack.
	• Unlike EtherCAT, RT PROFINET does not require specialized hardware to process and forward the Ethernet frames.
	• It provides support for TCP/IP and higher protocols (e.g., HTTP), making it very flexible and versatile.
		• The following basic requirements must be hold:
		  ○ QoS support.
		  ○ LLDP support.
		  ○ L2 multicast.
		• More advanced requirements must also be hold in a PROFINET network with CC-B devices:
		  ○ SNMP support.
		  ○ Switch embedding PROFINET IO device functions.

3.2 TSN-based solutions for deterministic 5G communication in industrial automation

Factories of the future will require highly reliable communication networks capable of supporting the stringent latency, bandwidth, and reliability requirements of demanding industrial applications. Moreover, the communication networks that will support Industry 4.0 must be able to adapt their operation to the changing connectivity and data transfer needs of industrial production systems. 5G networks and their future evolution have the necessary capabilities to meet these requirements. However, the use of 5G in industrial environments requires its effective and efficient integration with industrial wired networks currently implemented in factories to meet the strict reliability and resilience requirements of industrial applications. Time-Sensitive Networking (TSN) is becoming the standard wired technology for Industry 4.0 environments. TSN is based on Ethernet and provides unprecedented deterministic service levels with perfectly bounded latencies. However, TSN does not have the flexibility and reconfiguration capabilities required by the factories of the future. On the other hand, 5G is highly reconfigurable and flexible but cannot efficiently and scalable support deterministic services. Both networks, therefore, have complementary capabilities and the integration of the industrial 5G and TSN networks is key to support the Industry 4.0 paradigm.

3GPP standards already define the framework for the efficient integration of TSN and 5G networks. The 5G network is integrated into the TSN network as a logical TSN bridge of the TSN network (see

Project funded by Horizon Europe, Grant Agreement #101057083

Figure 19). A TSN bridge is an Ethernet switch that receives and transmits TSN frames based on a scheduling. In the integrated network shown in Figure 19, the communication path between two end-devices (for example, a sensor and an actuator) is established through one or more TSN bridges and the 5G logical bridge. In this context, a critical aspect to meet the end-to-end latency requirements of the industrial applications is the coordination of the scheduling of both networks. In the integrated 5G-TSN network, the Central Network Configuration (CNC) has information about the whole network topology and capabilities of the bridges. Based on this information, the CNC decides the communication path and scheduling for each TSN traffic flow. The scheduling decision establishes the arrival and departure time of the packets of each TSN flow in each (TSN and 5G logical) bridge. The scheduling of the packets of the TSN flows in the 5G network must be done in a way that these arrival and departure times calculated by TSN are satisfied, and finally guarantee the end-to-end latency requirements of TSN traffic. In Zero-Swarm, we propose a novel scheduling scheme for 5G integrated into a TSN network that aims to meet the latency requirements of the different TSN flows. The proposed scheme is based on Configured Grant (CG) and exploits the information provided by TSN about the characteristics of the TSN traffic to coordinate its decision with the scheduling of TSN. The final objective of the proposed CG scheduling scheme is to satisfy the end-to-end latency and deterministic requirements of industrial TSN traffic.

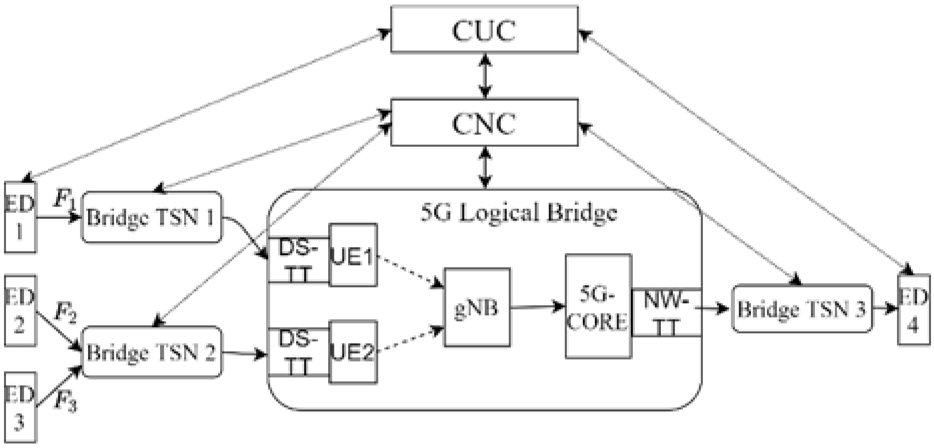

Figure 19. 5G-TSN integration model: 5G is integrated in the TSN network as a logical TSN bridge

3.2.1 5G-TSN integration

TSN is a set of open standards developed by the IEEE 802.1. TSN has been designed to provide communications with a very high level of determinism: TSN guarantees the delivery of the data in a guaranteed time window [[7]]. This is achieved thanks to the introduction of several technical features, such as strict time synchronization (IEEE 802.1AS) and a traffic scheduler (IEEE 802.1Qbv) that reserves specific time slots for the transmission of high-priority traffic. 5G networks have the capacity to support low latency and highly reliable communications. 5G can complement TSN to provide the level of determinism and flexibility demanded by Industry 4.0. 3GPP standards define the framework for the integration of 5G in TSN networks. A TSN network is integrated by end devices (ED), which are the source and destination of TSN flows, and bridges that are interconnected using standard Ethernet links. In the integration model defined by the 3GPP, 5G appears as a (logical) bridge of the TSN network, as shown in Figure 19. 5G includes TSN translator (TT) functionality at the edges of the network: at the users (UEs) referred to device side TT or DS-TT and at the 5G Core Network referred to network side TT or NS-TT. The TTs are the interconnection points between 5G and TSN. The TTs allow 5G to inform the TSN network about its capabilities and current status, as well as to receive, interpret and apply TSN configuration commands from the TSN network. The TTs provide TSN bridge ingress and egress port operations. For example, TTs are responsible for holding and forwarding the TSN flows to minimize latency variation.

The integrated 5G-TSN network considers a centralized management model. The TSN network incorporates the Centralized User Configuration (CUC) and CNC functions. The CUC receives information about the communication requirements for each TSN flow from the end devices. The CUC informs the CNC about these requests. The CNC collects information about the capabilities and current status of all the (TSN and 5G logical) bridges in the network [[8]][[9]]. For example, bridges inform about the minimum and maximum supported bridge delays, which are the minimum and maximum times respectively that a packet needs to be forwarded from an ingress port to an egress port of the bridge. TSN bridges also inform of the propagation delay or link delay, which is the time a packet takes to travel through the links that interconnect the nodes (end devices and bridges) from the specified port of the station to the neighboring port on a different station. The CNC then decides the path and schedule of all TSN flows using IEEE 802.1Q [[10]]. The CNC finally configures the (TSN and 5G logical) bridges according to the scheduling decision in order to guarantee the end-to-end requirements of the TSN flows. The CNC applies the time-aware shaper (TAS) defined in IEEE 802.1Qbv that establishes the arrival and departure time of each TSN flow at the ingress and egress ports of each TSN and 5G logical bridge, respectively.

5G translates the information about TSN QoS (Quality of Service) requirements received from the CNC to a 5G QoS Identifier (5QI) for each TSN flow. A 5QI defines a set of 5G QoS characteristics that describe the packet forwarding treatment that a QoS flow receives in the 5G network (between the UE and the last node in the 5G core network). A 5QI includes the priority level of a flow for the scheduling of resources, the packet error rate (PER), the maximum data burst volume (MDBV), and the packet delay budget (PDB), which indicates the maximum latency a packet can experience in the 5G network, among other QoS characteristics. 3GPP standards do not specify how TSN QoS requirements are mapped to 5G QoS requirements. Some proposals are found in [[11]]-[[14]]. In addition to the 5QI, 5G determines Time Sensitive Communication Assistance Information (TSCAI). The TSCAI describes the TSN traffic characteristics: burst or packet arrival time with reference to the ingress port, periodicity, flow direction (uplink or downlink), and survival time[1] [[9]]. The TSCAI and the 5QI may be used by the gNB (new generation Node B) to decide the scheduling of radio resources for the TSN traffic in order to meet the end-to-end requirements of the TSN flows.

3.2.2 State of the art of scheduling techniques for 5G networks integrated with TSN networks

Some current research works have studied scheduling schemes for 5G networks integrated into TSN networks as a logical bridge. For example, [[15]] proposed a dynamic scheduling scheme for deterministic traffic. The proposed scheme is based on optimization techniques and allocates resources dynamically for each TSN packet to improve radio resource efficiency while guaranteeing the latency requirements of the TSN traffic. With dynamic scheduling, end users need to request resources and/or receive a grant before transmitting a packet. This exchange of signaling can increase the latency of communication. Configured Grant (CG) in uplink (UL) and Semi-Persistent Scheduling (SPS) in downlink (DL) allocate radio resources periodically for each TSN flow based on the periodicity of the packets. Configured Grant or Semi-Persistent Scheduling are the most proper scheduling (in UL and DL, respectively) for this type of traffic based on its low latency requirements and periodicity. An SPS scheme has been studied in [[16]] for DL Time Sensitive Communications (TSC) flows. The authors in [10] proposed the adaptation of the modulation and coding scheme (MCS) used for the transmission of each packet based on past channel conditions. [[17]] proposed a predictive multi-priority scheduling

[1] Survival time is the time period an application can survive without receiving any burst or packet as described in [8].

Project funded by Horizon Europe, Grant Agreement #101057083

mechanism based on SPS for the 5G-TSN network where periodic and aperiodic data is considered. According to [[18]], CG or SPS are the most proper scheduling (in UL and DL, respectively) for TSC traffic demanding low latency requirements. The allocation of periodic resources to TSN flows with different periodicity may result in conflicts: the resources allocated to different TSN flows may overlap after several periods. This is illustrated in Figure 20. The example in Figure 20 shows two TSN flows, with periodicity 3 and 5 respectively. Although the radio resources allocated for the transmission of the first packet of each TSN flow are different, some of the resources allocated for the transmission of the third and fifth packets of F1 and F2 respectively overlap in frequency and time due to the different packet periodicities of F1 and F2. This is a critical issue that, to the best of the author's knowledge, has not been previously addressed. In this context, this document progresses the state of the art by proposing a novel scheduling scheme for 5G integrated into TSN networks that addresses this problem. The main objective of the proposed scheduling scheme is to meet the latency requirements of the different TSN flows based on the arrival and departure times calculated by the TSN network. To this end, the proposed scheduling scheme uses the information provided by the TSN network about the characteristics of the TSN flows (packet sizes, periodicity, etc.). The proposed scheduling scheme configures several grants for each TSN flow in order to avoid radio resource allocation conflicts among different TSN flows. This study demonstrates that the proposed scheduling scheme considerably increases the number of TSN flows that can be satisfactorily served in the integrated TSN-5G network.

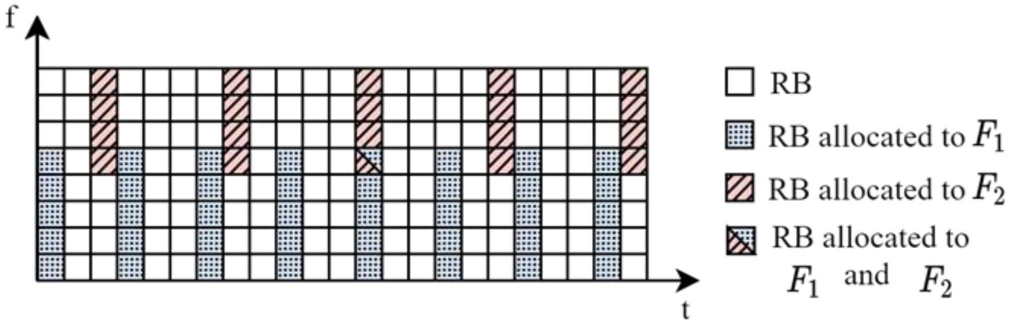

Figure 20. Example of a radio resource allocation conflict where the same radio resources are allocated to more than one TSN flow when radio resources are allocated considering the packet periodicity

3.2.3 5G CG scheduling for the support of TSN traffic

We propose a scheduling scheme for a 5G network integrated into a TSN network to support the stringent communication requirements of industrial applications. The main objective of the proposed scheduling scheme is to guarantee the 5G latency requirements for each TSN flow and ensure that packets are received at the egress port of the 5G logical bridge before the departure time indicated by the TSN network. To this end, the proposed scheduling scheme uses the information provided by the TSN network (5QI and TSCAI) to decide the radio resource allocation solution that better satisfies the latency requirements of all TSN flows. A radio resource is composed of one Resource Block (RB) of 12 subcarriers in the frequency domain and one Orthogonal Frequency Division Multiplexing (OFDM) symbol in time domain. The scheduling proposal is based on CG, and it allocates radio resources periodically for each TSN flow. The proposed scheduling scheme configures several grants for each TSN flow, if needed, to ensure that radio resources are not assigned to more than one TSN flow simultaneously.

Let us consider that NF TSN flows must be transmitted through the 5G virtual bridge toward the end devices. For each TSN flow F_i (with i=1,…, N_F), packets of size s_i arrive with a periodicity p_i; s_i is given by the MDBV of the 5QI and p_i in the TSCAI. s_i and p_i can be different for each TSN flow. We consider that all the TSN traffic that arrives to the 5G network must be transmitted in UL (flow direction is represented by f_i). Each flow demands d_i radio resources calculated as a function of s_i and the MCS to use for the packet transmission. Packets in a TSN flow F_i are referred to as $pkt_{i,j}$ where j indicates the number of the packet in the flow. The arrival and departure time instants of a packet $pkt_{i,j}$ are given by $A_{i,j}=A_i+jp_i$ and $D_{i,j}=A_{i,j}+l_{5Gi}$, respectively, where A_i is the burst arrival time (BAT) indicated in the TSCAI and is the latency requirement for packets in the TSN flow F_i indicated by the PDB of the 5QI. A TSN flow F_i is then defined as $F_i=\{A_i, , d_i, p_i, f_i\}$. The proposed scheduling scheme exploits this information to decide the radio resources that must be allocated to each TSN flow.

The TSN flows supported by the 5G network transmit periodic packets. Although the periodicity p_i of each TSN flow can be different, it is possible to find a pattern of packet arrival considering all TSN flows that repeat periodically. The periodicity of this packet arrival pattern is referred to as hyperperiod or HP. Figure 21 shows an example with three TSN flows with different periodicities (30, 45 and 90 ms). It is possible to see in that there is a packet arrival pattern that repeats every 90 ms when considering the set of all TSN flows. The scheduling scheme identifies HP and allocates radio resources for each packet within HP. The allocated radio resources for each TSN flow are pre-assigned periodically with a periodicity HP.

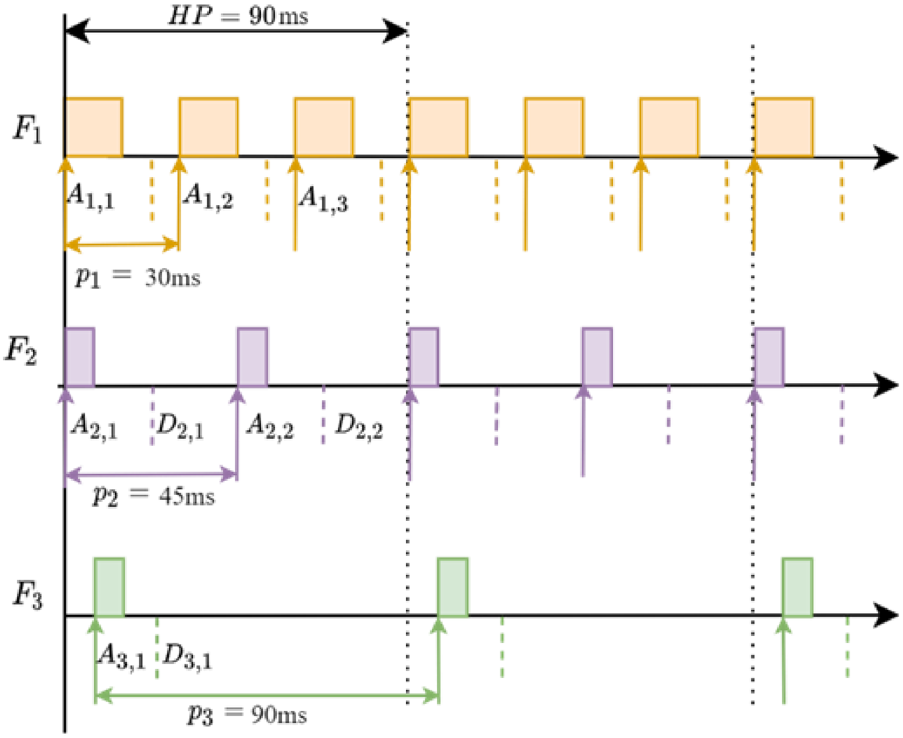

Figure 21. Hiperperiod for an example case with 3 TSN flows with different periodicities

The operation of the proposed scheduling scheme is presented in Algorithm I and Algorithm II. The scheduling scheme first calculates *HP*. *HP* is calculated as the least common multiple (lcm) of the periodicity of the packets transmitted in each TSN flow, i.e., HP = lcm(p_i), with i [1, N_F] (line 2 in Algorithm I). The packets of all the TSN flows within the time window given by *HP* are divided into different groups or blocks. Each block contains the packets whose transmission could overlap in time considering their arrival and departure times. Two packets $pkt_{i,j}$ and $pkt_{m,n}$ overlap if $A_{i,j} \le A_{m,n} \le D_{i,j}$ or $A_{i,j} \le D_{m,n} \le D_{i,j}$ (see Algorithm II). In the example of Figure 21, two blocks are created: the first block B_1 contains $pkt_{1,1}$, $pkt_{2,1}$ and $pkt_{3,1}$, while the second block B_2 contains $pkt_{1,2}$, $pkt_{2,2}$ and $pkt_{1,3}$. The proposed scheme addresses the radio resource allocation process separately for each block B_z, with z=1,2,... First, the scheduling scheme calculates the number of symbols and RBs that should be allocated for the transmission of each packet to satisfy d_i. To this end, the scheduling scheme is based on the Sym-OFDMA (Orthogonal Frequency Division Multiple Access) scheduling policy presented in [[19]]. Sym-OFDMA has been selected because it reduces the transmission latency compared to other scheduling policies. Following Sym-OFDMA, the scheduling scheme tries to minimize the number of symbols assigned for the transmission of each packet. In this context, the scheduling scheme will allocate d_i^R= d_i RBs in one symbol (d_i^S=1) when d_i is lower than the number of RBs available in the bandwidth (R_{BW}), i.e. $d_i \leq R_{BW}$,. If $d_i > R_{BW}$, the scheduling scheme will allocate $d_i^R = R_{BW}$ RBs in $d_i^S = \lceil d_i / R_{BW} \rceil$ symbols. When d_i^S and d_i^R are known, the scheduling scheme starts an iterative process. At each iteration, the scheduling scheme tries to find a radio resource allocation solution that satisfies the latency and departure times for all packets. The scheduling scheme starts a new iteration if it cannot find adequate radio resources to satisfy the departure time for a packet.

ALGORITHM I: *SCHEDULING*

```
1.  Input: F_i ∀ i ∈ [1, N_F], t_sym
2.  HP = lcm{p_i} with i ∈ [1, N_F]
3.  Create packet blocks B_z, with z=1, 2, ... (Algorithm II)
4.  Initialize BF_z ≠ Ø
5.  For z=1 to N_B
6.    Calculates d_i^S and d_i^R for all packets in B_z
7.    N_pkt = number of packets in B_z
8.    pkt_i,j = packet with lower l_5G in B_z
9.    B_z = B_z-{pkt_i,j}
10.   iter=0, BF_z = Ø
11.   For iter = 0 to ∞
12.     While B_z ≠ Ø
13.         s = first OFDM symbol after A_i,j + t_UE,tx
14.         While s + d_i^S-1 <= D_i,j
15.           If there are d_i^R unallocated RBs in symbols s until s+ d_i^S-1
16.             pkt_i,j receives d_i^R RBs in symbols s until s+ d_i^S-1
17.             First allocated symbol: s_i,j = s
18.             If s + d_i^S-1 == D_i,j
19.                 BF_z = BF_z + {pkt_i,j}
20.             EndIf
21.           Else
22.             s++
23.           EndIf
24.         EndWhile
25.         If s + d_i^S-1 > D_i,j
26.           BF_z = BF_z + {pkt_i,j}
27.           If BF_z + B_z == N_pkt
28.             Goto 41
29.           Else
30.            Reinitiate B_z
31.            B_z = B_z - BF_z
32.            Free radio resources allocated to packets in B_z
33.            Goto line 11
34.           EndIf
35.         Else
36.           pkt_i,j = packet with lower l_5G in B_z
37.           B_z = B_z-{pkt_i,j}
38.         EndIf
39.     EndWhile
40.   EndFor
41. End For
```

Project funded by Horizon Europe, Grant Agreement #101057083

```
ALGORITHM II: BLOCK CREATION
1.  Input: F_i ∀ i ∈ [1, N_F]
2.  I={A_i,j} ∀i ∈ [1, N_F], ∀j ∈ [0, HP/p_i]
3.  z=1, B_z=Ø
4.  While 1
5.    {i, j} = arg_{m,n} {min_{m,n}{I}} ∀m ∈ [1, N_F], ∀n ∈ [1, HP/p_m]
6.    I = I – {A_i,j}
7.    If B_z==Ø
8.      D_aux = D_i,j, A_aux = A_i,j
9.      B_z = {pkt_i,j}
10.   Else
11.     If A_i,j <= D_aux
12.       B_z = B_z ∪ {pkt_i,j}
13.       If D_i,j > D_aux
14.         D_aux = D_i,j
15.       End If
16.     Else
17.       z=z+1, B_z=Ø
18.       Go to 7
19.     End If
20.   End If
21.   If I==Ø
22.     Go to 24
23.   End If
24. End While
```

The iterative process will finalize when one of the two following conditions are met: all packets in B_z have allocated resources, or it is not possible to find a feasible solution to the problem. At each iteration, the proposed scheduling scheme serves packets considering their latency requirements l_{5Gi}: packets with lower l_{5Gi} are served first (line 8 in Algorithm I). The scheduling scheme allocates to each packet $pkt_{i,j}$ d_i^R RBs in the first consecutive symbols with at least unallocated RBs after the packet is generated ($A_{i,j}+t_{UE,tx}$, where $t_{UE,tx}$ represents the processing time in the transmitter to generate the packet) and before the departure time $D_{i,j}$ (lines 15-21 in Algorithm I). If a packet receives radio resources just before its departure time $D_{i,j}$, the packet will maintain the allocated radio resources although a new iteration is performed. To this end, it is included in a set BF_z. BF_z includes the packets that cannot change their allocated radio resources in potential next iterations (see lines 18-20 in Algorithm I). If it is not possible to find the required RBs and OFDM symbols for a packet, the scheduling scheme will start a new iteration. Before that, B_z is reinitialized but excluding packets in BF_z (lines 25-35 in Algorithm I). The first packet served in the new iteration will be the packet that was not possible to assign radio resources in the previous iteration. This packet is also included in BF_z (line 26 in Algorithm I). The scheduling process finalizes when all packets have allocated resources or when it is not possible to find a feasible solution to the problem (lines 27-29).

3.2.4 Evaluation scenario

We consider an evaluation scenario where a 5G network is integrated into a TSN network as a logical bridge. The integrated 5G-TSN network provides connectivity to an industrial plant where there is implemented a closed-loop supervisory controller application [[20]]. This application integrates a Programmable Logic Controller (PLC) which receives monitoring data from sensors (S1, S2, …, S_{N_F}) that creates a total of NF TSN flows. Based on the received data, the PLC sends a command to an actuator

(A). Figure 22 shows the evaluated scenario. The PLC, sensors, and actuator are the static end devices that exchange data in the network. The end devices are connected through several bridges, including the 5G logical bridge. The evaluation of the proposed scheduler has been carried out using Matlab.

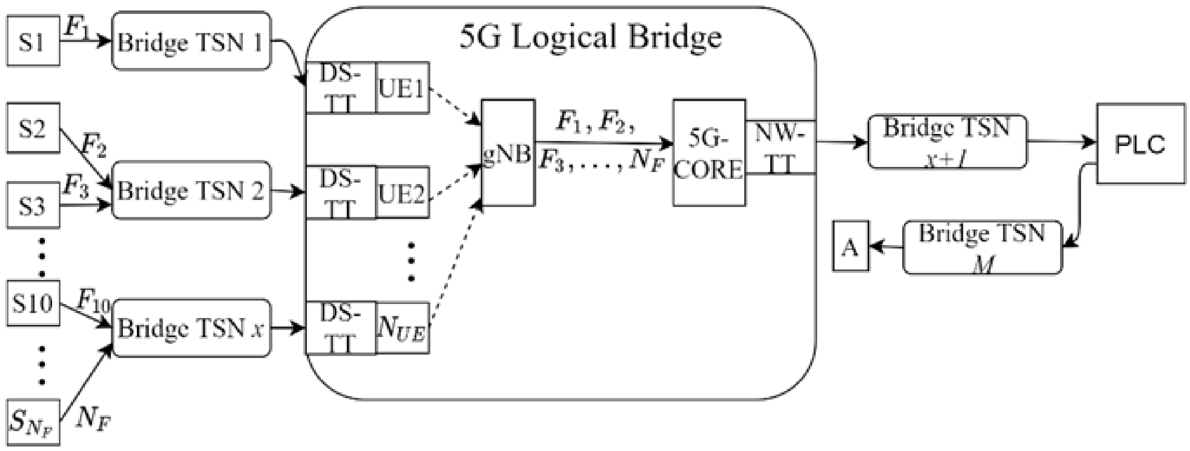

Figure 22. Evaluated scenario

The sensors send the sensed data to the PLC. This results in N_F TSN flows between the sensors and the PLC that are transmitted through the 5G logical bridge (we consider N_F equal to 30). Each TSN flow $F_i$is defined by the size of the transmitted packets (s_i), the periodicity of the packets (p_i) and the demanded end-to-end latency (l_{e2e_i}) between the sensor and the PLC. Based on [[20]], l_{e2e_i}can take values between 4ms and 20 ms. We consider that the periodicity p_i is equal to the l_{e2e_i} requirement (this means that each packet needs to be received before the next packet is generated). TSN flows randomly select the periodicity p_i between p_1, p_2 or p_3, and p_j, where $j \in \{1,2,3\}$, is randomly selected between 4 and 20 ms. The packet size for each TSN flow F_i can take a value between 40 and 250 bytes. The characteristics and requirements that need to be satisfied by the 5G logical bridge of a TSN flow is indicated to the 5G network using the TSCAI and 5QI information. For each TSN flow, the TSN network informs the 5G network about the following information: the periodicity p_i and packet size s_i of the packets, the flow direction f_i, the arrival time at the ingress port of the 5G logical bridge, and the maximum latency requirement of the TSN flow inside the 5G logical bridge l_{5G_i}. In the evaluation scenario, all TSN flows in the 5G network are transmitted in uplink, i.e., from the UEs to the gNB. The arrival time A_i and the 5G latency requirement is calculated by the CNC that has information about the TSN network topology and is in charge of the scheduling of the TSN flows in the TSN network to satisfy the end-to-end latency requirements. The arrival time A_i of the first packet of a TSN flow is equal to the latency that the packet experiences from the moment it is generated in the sensor until it is received in the 5G logical bridge. Then, it depends on the application processing time in the sensor l_{sensor_i}, the number and the processing time of TSN bridges that there are in the path between the sensor that generates the data and the 5G logical bridge $l_{TSN_{b,i}}$ with $b \in [a,..., x]$, the propagation time it takes a packet to travel through the links that interconnect the nodes (end devices and bridges) l_{links_i}, and the processing time $t_{DS\text{-}TT}$ at the DS-TT translator. The is given by the following expression:

$$A_i = l_{sensor_i} + l_{links_i} + l_{TSN_{a,i}} + \ldots + l_{TSN_{x,i}} + t_{DS\text{-}TT} \quad (1)$$

$l_{TSN_{b,i}}$ represents the time that packets of a TSN flow i spend in a TSN bridge b. $l_{TSN_{b,i}}$ is null if TSN bridge b is not in the path between the sensor that generates the data and the 5G logical bridge. Otherwise, $l_{TSN_{b,i}}$ considers processing and queuing delays in the TSN bridge (independent delay $t_{indep_{b,i}}$) that vary depending on the traffic load [[9]], and the transmission delay that depends on the packet size and the bit rate used by the TSN bridge to transmit the data to the next node in the path (dependent delay $t_{dep_{b,i}}$). To calculate the arrival time, we consider that all sensors (30 sensors) connect to a TSN bridge that connects to the 5G logical bridge, and all sensors transmit packets of 250 bytes. Due to short distances, l_{links_i} is considered negligible. We consider DS-TT and NW-TT processing times equal to 0.05 ms and 0, respectively [[21]]. We assume a data bit rate equal to 100 Mbps, and the minimum independent delay $t_{indep_{b,i}}$ is equal to the TSN bridge processing delay, which is 1.5 µs as presented in [[22]]. Then, for the minimum arrival time, we assume that a flow is transmitted through the TSN bridge directly to the 5G logical bridge as soon as it arrives to the ingress port of the TSN bridge. We have considered that the sensor processing time is 3 µs [[23]], then, the minimum arrival time to 5G logical bridge for a packet of 250 bytes is $3\mu s+\left(\frac{250\cdot 8bits}{100Mbps}+1.5\mu s\right)+50\mu s \cong 75\mu s$. In the case of 700 µs which represents the maximum arrival time, we consider the last flow to be transmitted from TSN bridge to the 5G logical bridge arrives at $3\mu s+30\cdot\left(\frac{250\cdot 8bits}{100Mbps}+1.5\mu s\right)+50\mu s\cong 700\mu s$ to the ingress port of 5G logical bridge. This results in arrival times between 75 µs and 700 µs.

We calculate the latency that must be satisfied in the 5G network for each TSN flow F_i as:

$$l_{5G_i} \leq l_{e2e_i} - A_i - l_{links_i} - \sum_{b=x+1}^{M} l_{TSN_{b,i}} - t_{NW\text{-}TT} - l_{PLC_i} \quad (2)$$

In (2), l_{PLC_i} represents the application processing time in the PLC, $t_{NW\text{-}TT}$ is the processing time at the NW-TT translator, $l_{TSN_{b,i}}$ represents the time that packets of a TSN flow *i* spend in a TSN bridge *b* with $b \in [x+1,..., M]$, *M* is the number of TSN bridges in the TSN flow and l_{links_i} is the propagation time it takes a packet to travel through the links that interconnect the nodes from the 5G logical egress port until the PLC ingress port. To calculate l_{5G_i}, we assume l_{PLC_i} equal to 1007 µs as we consider a scan time of analogue input data and the execution time of Proportional Integral Derivative (PID) control in the PLC of 1ms [[24]] and 7µs [[25]], respectively. We have assumed that the scan and processing times are minimal. This data is an approximation since it will vary with the application.

The 5G network deployed in the industrial plant implements a single cell with a bandwidth of 20 MHz. We consider the use of a 30 kHz sub-carrier spacing (SCS) as recommended in [[26]] for industrial environments. The cell operates in TDD mode [[27]], and each slot reserves 2 OFDM symbols for the transmission of UL and DL control messages and 12 OFDM symbols for the transmission of UL data. The UEs are placed in locations that guarantee Line of Sight conditions with the gNB. Based on that, we consider that all packets are transmitted using the MCS 12 in MCS table 1 [[28]] which guarantees a good compromise between robustness and spectral efficiency (modulation order Q_m equal to 4 and coding rate equal to 434/1024). The number of radio resources to transmit a packet of size s_i can be then calculated as:

$$d_i = \left\lceil \frac{(tbs_i(s_i+s_{header})+s_{CRC})\cdot 8}{R\cdot Q_m \cdot N_{sc,RB}} \right\rceil \quad (3)$$

In (3), s_{header} represents the size in bits of the IPv4 header added to the data packet, and s_{CRC} is the size in bytes of the cyclic redundancy check (CRC) code added at the end of the packet. $tbs(s_i+s_{header})$ represents the smallest transport block size from the available values given in [[29]] that can be used to transmit s_i+s_{header} bits, and $N_{sc,RB}$ is the number of subcarriers in a resource block ($N_{sc,RB}$=12).

3.2.5 3.2.5. Performance evaluation

This section presents the evaluation results achieved with the proposed scheduling scheme proposed for a 5G network integrated into a TSN network to provide connectivity in industrial scenarios. To this end, we compare its performance to that achieved with a commonly used CG scheduling scheme. The reference scheme allocates radio resources periodically for each TSN flow F_i with a periodicity p_i. Each TSN flow F_i receives the number of radio resources necessary to transmit packets of size s_i according to (3). For a fair comparison with the proposed scheduling scheme, the reference scheme serves TSN flows based on their 5G latency requirements: TSN flows with lower l_{5G_i} values are served first. Since the TSN flows can have different periodicities, several TSN flows could receive the same radio resources after some periods.

Due to the limited availability of radio resources and the high resource demands, it can be difficult for scheduling schemes to find a valid solution that can satisfy all requests. In this context, we first analyze the capabilities of the proposed and reference scheduling schemes to achieve a solution that allocates the requested radio resources for all the TSN flows. Figure 23 shows the percentage of evaluated scheduling problems for which the proposed and reference schemes satisfy the radio resource request for all the TSN flows. The results in Figure 23 are shown for different numbers of TSN flows demanding resources in the 5G network (between 10 and 30). Figure 23 (lines with circles) considers that the end-to-end latency (l_{e2e_i}) demanded by each TSN flow F_i is equal to its packet periodicity p_i, and Figure 23 (lines with starts) considers that all TSN flows demand a l_{e2e_i} equal to 4 ms. The results in Figure 23 show that when the latency requirements are more relaxed, both the proposal and the reference scheme obtained solutions that satisfy all the requests. When the latency requirements are more stringent (l_{e2e_i}=4 ms for all the TSN flows) and the number of TSN flows requesting resources increases to 25, the reference scheme fails to find solutions satisfying all the requests in 9% of the cases. In this scenario, the proposal achieves solutions satisfying all the requests in the 98% of the cases. When the number of TSN flows increases to 30, the proposal is able to satisfy all the requests in the 33% of the cases, while the reference scheme only in the 17%.

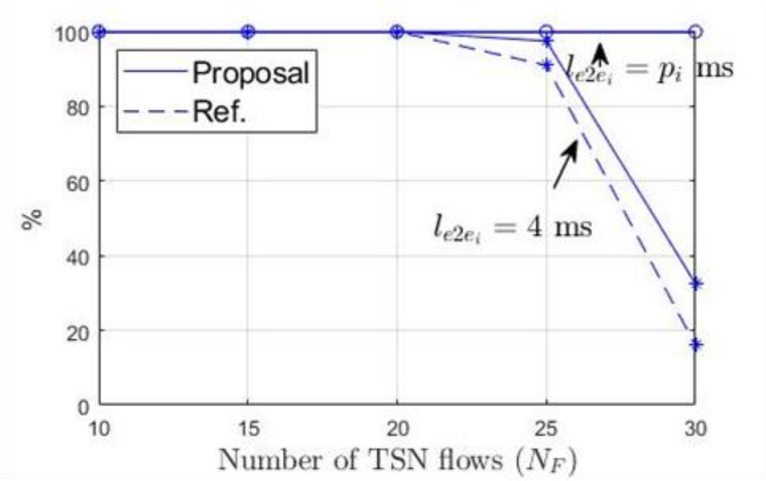

Figure 23. Percentage of times the proposed and reference scheduling schemes satisfy the radio resource request for all packets of all TSN flows as a function of the number of TSN flows

The solutions achieved with the reference scheme ensure that the radio resources allocated to the TSN flows do not overlap in time and frequency simultaneously for the first packet in each TSN flow. However, this is not guaranteed for the following packets in the flows, and several TSN flows can receive the same radio resources after several periods due to the different periodicities of each TSN flow (Figure 20). This is a radio resource allocation conflict that can ultimately result in packet collisions. Figure 24 shows the percentage of times that the reference scheme achieves a solution serving all requests, but some radio resources are allocated to more than one TSN flows. The results in Figure 24 show that the number of times that the solution achieved with the reference scheme results in a radio resource allocation conflict increases with the number of TSN flows demanding resources. It is important to note that the percentage of times that there is a resource allocation conflict with the reference scheme is higher than 74% when the number of TSN flows demanding resources is equal to or higher than 15.

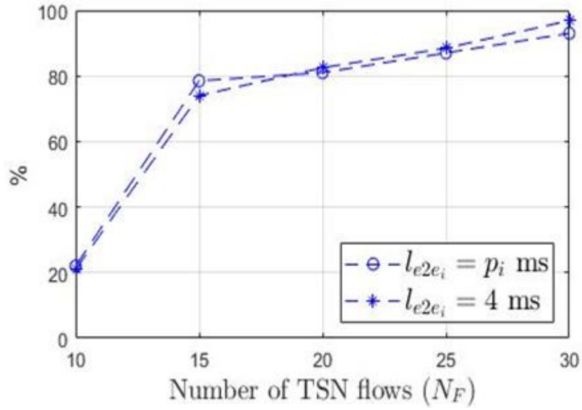

Figure 24. Percentage of times that the reference scheme achieves a solution serving all requests, but some radio resources are allocated to more than one TSN flows as a function of the number of TSN flows

Figure 25 depicts the average of the latency experienced by the packets of each TSN flow when a different number of TSN flows are transmitted in the 5G network considering that the end-to-end latency (l_{e2e_i}) demanded by each TSN flow F_i is equal to its packet periodicity p_i. The results show that the proposed scheduling scheme reduces the average latency experienced by the transmitted packets by more than 40% when the number of TSN flows is equal to or higher than 25. This is as a result of the higher flexibility of the proposed scheduling scheme. Thanks to the flexibility of the proposed scheduling scheme, the proposal assigns for each packet of a TSN flow the radio resources that minimize the latency based on the availability of unassigned resources at each moment. This results in a reduction of the latency experienced in the 5G network.

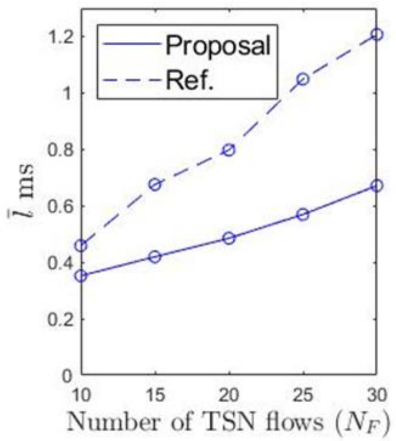

Figure 25. Average latency experienced with the proposed and the reference scheduling schemes as a function of the number of TSN flows when is equal to the periodicity

3.2.6 Summary

As a summary, in this section it has been presented the results achieved with a novel Configured Grant scheduling scheme for 5G networks integrated with TSN networks proposed within the Zero-Swarm project. In this study, we have considered the 5G-TSN integration model proposed by 3GPP where 5G is integrated into the TSN network as a virtual bridge. The designed scheduling scheme aims to coordinate its scheduling decision with the global scheduling performed in TSN to meet efficiently and effectively the end-to-end latency and determinism requirements of TSN traffic. To this end, the proposed scheduling scheme exploits the information provided by TSN about the characteristics of the TSN traffic (periodicity, packet size, etc.) to satisfy the maximum latency established by the TSN scheduling decision.

The proposed scheme tries to avoid scheduling conflicts that can arise with common CG scheduling schemes that allocate radio resources periodically for each TSN flow based on the periodicity of the packets. When TSN flows can different periodicities, this common CG schemes may result in the allocation of the same radio resources for the transmission of packets of different TSN flows. The proposed CG scheduling scheme solves this issue by configuring several grants for each TSN flow in order to avoid radio resource allocation conflicts among different TSN flows. The results achieved have demonstrated that the proposed scheduling scheme considerably increases the number of TSN flows that can be satisfactorily served in the integrated 5G-TSN network compared to common CG schemes. Within Zero-Swarm, we will continue working on the enhancement of the proposed scheduling scheme to ensure that it can be implemented in current 5G networks. In addition, we aim to analyze how much current solution could be improved using optimization techniques.

3.3 Other solutions for deterministic 5G communication in industrial automation

3.3.1 Cross layer scheduling optimization based on RAN feedback

3GPP RAN can optimize its scheduling based on the TSCAI (e.g. Periodicity, burst arrival time) in Rel-16 5GS Bridge case and Rel-17 TSC case. Deficiencies are that RAN just receives the traffic flow periodicity and burst arrival times (BATs) but cannot influence them.

One of the potential problems especially considering low latency applications is that the provided times may not fit well with the TDD cycle used in the network.

This creates additional delay or delay variation to the traffic flows. This can be an issue for QoS Flows requiring PDB 5 ms or lower. One simplified example for DL in the worst case is when the burst arrives just after RAN scheduling decision is made, the additional buffering time before the next scheduling cycle could be up to 2 ms. On the other hand, with best case with perfect match between TDD pattern and TSC burst arrival time, this time is negligible.

Additionally, since the burst arrival times of different flows traversing in the same direction through the same RAN are not coordinated they can collide with each other creating an additional source for delay variation. Radio configuration may vary cell by cell and the UEs may experience cell changes due to mobility, nomadicity or changes in radio environment, that's why the periodicity coordination must be flexible enough. This applies to both UL and DL direction.

UL may be even more important as in TDD the UL time duration is typically shorter than the DL one. Thus, it would be beneficial if the 5GS could influence the transmission cycle and time of the extremely low latency periodic traffic flows in such a way that they can be well aligned with the RAN TDD cycle and with each other for transmission over the radio link. This would reduce delay and delay variation for the traffic flows and could also allow 5GS to serve more traffic flows with required QoS.

In order to align the arrival of the traffic burst and periodicity of the traffic bursts with the expected time interval between subsequent transmission opportunities over the air interface in each direction (i.e. DL or UL), NG-RAN can determine an adjusted Periodicity along with above specified BAT offset. NG-RAN may support the following feedback mechanisms:

- Proactive RAN feedback for adaptation of Burst Arrival Time and Periodicity: NG-RAN may provide a Burst Arrival Time offset and an adjusted Periodicity based on received Burst Arrival Time and Burst Arrival Time Window from the application.
- Reactive RAN feedback for Burst Arrival Time adaptation: NG-RAN may provide a Burst Arrival Time offset in reference to the current arrival time of the bursts experienced by RAN in DL and by UE in UL.

Besides that, for UE-to-NW-to-UE industrial communication illustrated in Figure 26, how burst arrival time offset is synchronized between two involved RANs needs to be investigated since the RAN provides feedback for UL stream and DL stream independently. Basically, the 5GC needs to coordinate the BAT offset respectively received for UL leg and DL leg for a UE-to-NW-to-UE stream to avoid multiple times negotiation for adjustment.

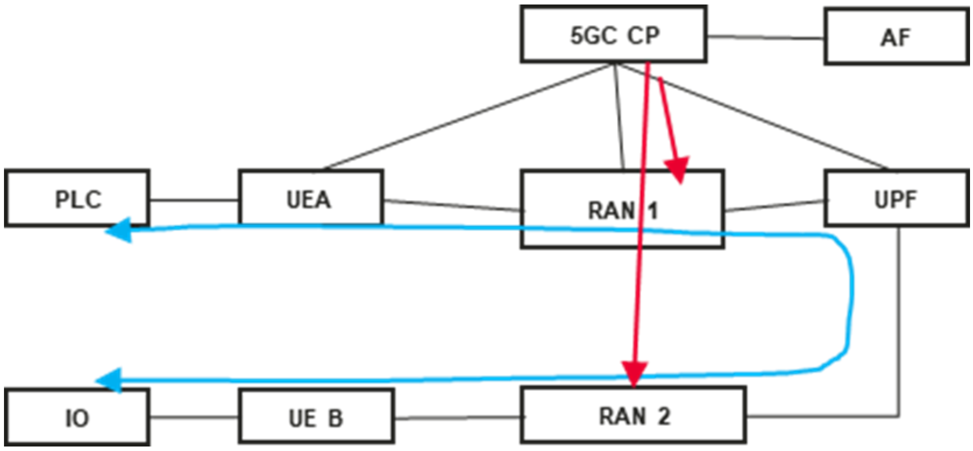

Figure 26 UE-network-UE communication by means of time synchronisation between two RANs

3.3.2 5G offering with higher cell capacity over lower latency and higher deterministic industrial connectivity

This is to support actual factory deployment with higher cell capacity while keeping industrial connectivity with lower latency and higher deterministic, for example QoS requirements: reliability 99.999%, deterministic delay 2-4ms and Air interface capacity: concurrent access of 500 to 1000 terminals per cell. To achieve higher cell capacity and link reliability, 5GS can further investigate on the following aspects:

- *User Plane Content Compression:* Ethernet frame has a minimum packet size of 64 Bytes, comprising an 18-Byte header and a payload of 46 Bytes. The packet payload of industrial applications is usually small and ranges from 10 bytes to 30 bytes, e.g. 20 bytes (Please see TS 22.104 [1]), while the rest of the payload is filled up with padding. In addition, there are a lot of duplicated packets in the industry scenario, the content (except the sequence number) in the packet payload does not change unless there is a status update, such as an alarm. In all these cases, it is valuable to study *how to avoid transmission of the unnecessary bits* (e.g. padding) over the air interface thus to improve the efficiency of the air interface. Basically, when sending the Ethernets frames over NR interface, the UE (UL) or the UPF (DL) identify and remove the unnecessary bits (e.g. padding) for the Ethernet frames according to the QoS rules or N4 rules; when receiving the Ethernet frames from NR interface, the UE (DL) or the UPF (UL) recover the unnecessary bits (e.g. padding) for the Ethernet frames according to the QoS rules or N4 rules before transferring the Ethernet frames to the downstream nodes. This procedure is illustrated in Figure 27.

UE2UE Survival time enhancement: Industry services usually have survival time which is the maximum interruption time that the factory service keep alive lock can accept, namely the time that an application consuming a communication service may continue without an anticipated message or data burst. 5GS supports provisioning survival time as one of the traffic characteristics for a stream either in terms of maximum number of messages (message is equivalent to all packets of a data burst) or in terms of time units.

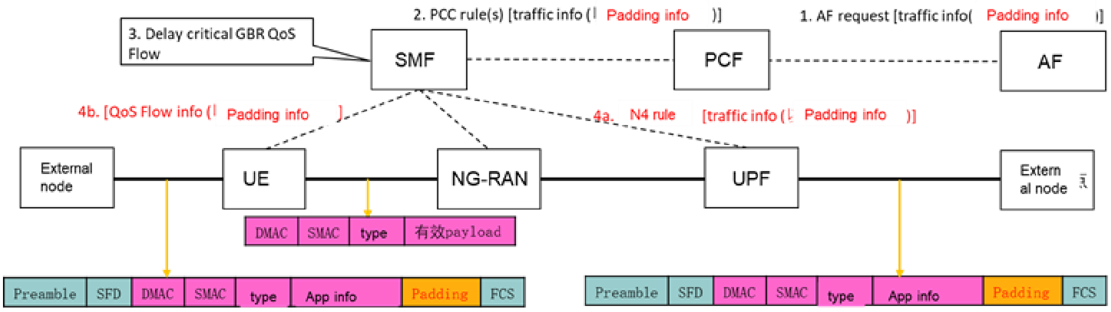

Figure 27 Padding procedure for Ethernet frames

Only a single data burst is expected within a single time period referred to as the periodicity. For UE-to-NW-to-UE industrial communication, how survival time is ensured still needs to be studied, e.g. how the RAN nodes coordinate the packet loss. If there is one packet loss for a specific stream at NG-RAN 1, then the NG-RAN 2 needs to be aware of the packet loss at NG-RAN 1 for that stream via control plane or user plane, thus to continue counting the packet loss towards the survival time in subsequent stream transmission and decide whether to increase the scheduling priority for the next packet of that stream.

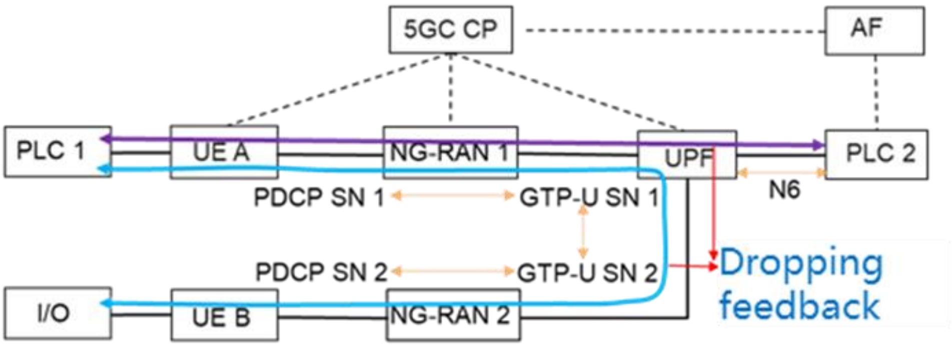

Figure 28 UE-network-UE communication by means of packet loss coordination between two RANs

3.3.3 Integration of 5G VN and legacy LAN for a wider 5G LAN-type service

In campus network deployment, there is a need to support interconnection between different sites and enterprise clouds as illustrated in Figure 29, for example bank branches, chain stores, factory areas, enterprise campuses, and electric grids. This corresponds to the leftover SA1 requirement in TS22.261, clause 6.24: The Ethernet transport service shall support the transport of Ethernet frames between a UE that an Ethernet device is connected to and an Ethernet network in DN (Data Network). Now the 5GLAN feature supports static connectivity between UPFs controlled by different SMF sets, but lacks the support of dynamic connectivity between UPFs controlled by different SMF sets and connectivity between UPFs and the enterprise cloud. To achieve the Integration of 5G VN and legacy LAN, the 5GS needs to support the following:

- Integration architecture: cross-site tunnel establishment and release, QoS provisioning and monitoring, intelligent traffic steering, exposure for reporting and configuration mapping
- Management of the DN side devices and the UE side devices as 5G VN group members
- Removal of the standard restrictions e.g. " Multicast group formation of selected members of a 5G VN for Ethernet type data communication is not described ", "No support of loop-free

Project funded by Horizon Europe, Grant Agreement #101057083

due to topology changes", "restriction for maximum of 16 VLAN tags or maximum of allowed MAC addresses" etc.

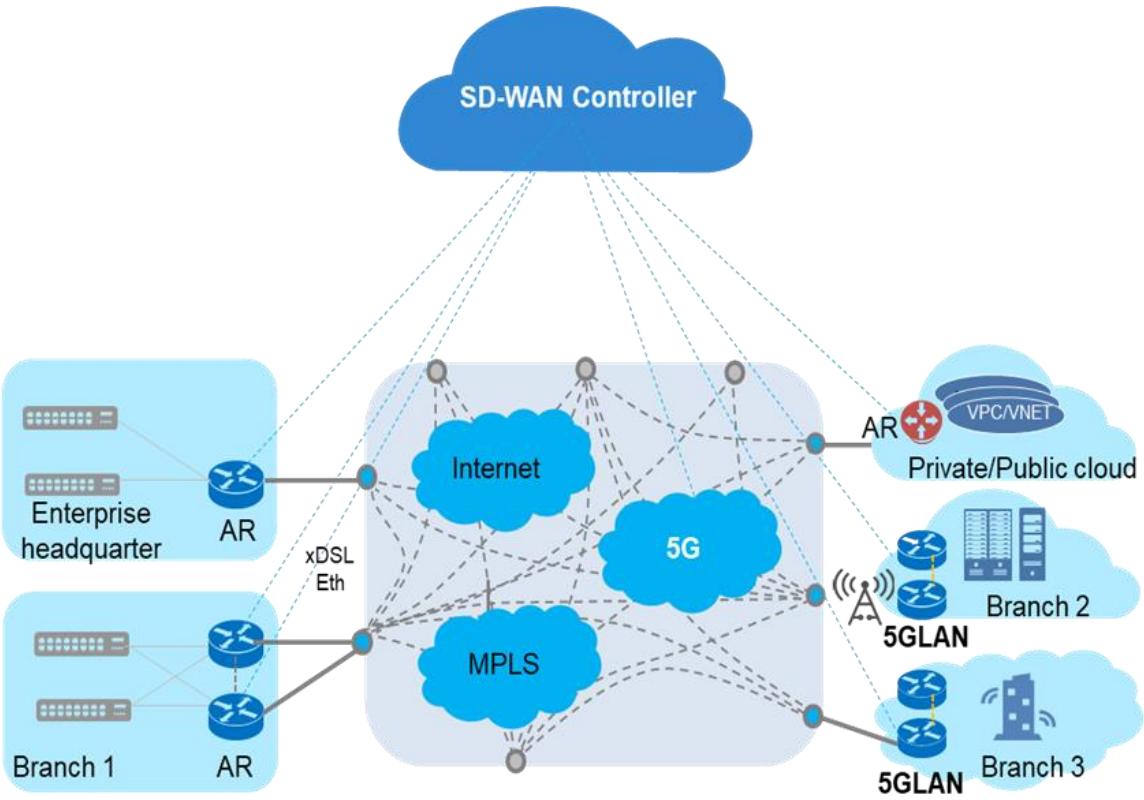

Figure 29 Support integration of 5G VN and legacy LAN.

Besides that, there is a need to further address Identified Gaps and Limitation for Exposure of 5G Capabilities in 5G-ACIA. The exposure of 5G capabilities for industrial applications is illustrated in Figure 30. The 5G-ACIA has identified gaps and limitations regarding the exposure of 5G capabilities for industrial applications. To enable easy and in-depth Interconnection with Verticals without having them to have in-depth knowledge on the underlying 5G technology, it is important to improve NEF exposure capability to address 5G-ACIA identified gaps and limitation [[30]]:

- Acknowledging an AF service request within indicated time
- VLAN tag assignment without AAA-Server
- UE-to-UE QoS with one NEF API
- Per QoS characteristic monitoring
- IP address assignment via NEF exposure

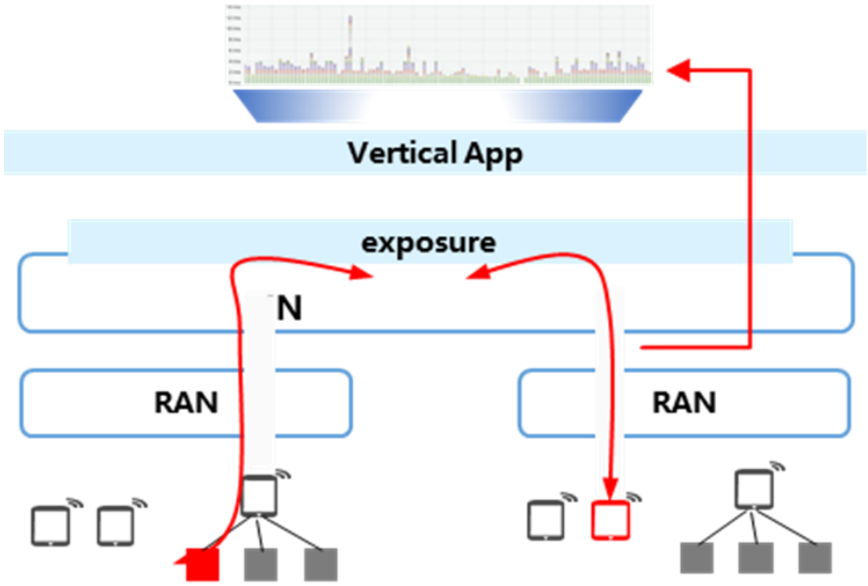

Figure 30 Exposure of 5G capabilities for industrial applications.

3.3.4 Support of redundant user plane paths based on multiple UEs at 3GPP layer to adapt to Ethernet industrial redundancy protocols

The reliability with a single UE suffers due to the radio channel quality, this becomes worse in the circumstance that has a lot of steels, e.g. warehouse, harbour, workshop, where the location of the work equipment varies much resulting poor connection stability. To break the limitation of reliability dependency on a single UE, the 5G system needs to consider the support of user plan redundancy with multiple UEs, thus to achieve reliability higher by using additional UE in an incremental manner. Now the URLLC feature supports redundant user plane paths based on multiple UEs per device, but the upper layer protocols are assumed to make use of the multiple user plane paths for high reliability, to lift this assumption to remove the restriction on a flexible deployment, there is a need to investigate how to support of replication and deduplication at redundant user plan path i.e. UE and UPF as illustrated in Figure 31. Besides, more technical issues need further study, e.g. forwarding packets with same MAC address to different redundant user plan paths, topology adaptation for industrial redundancy, etc.

5G native higher reliability with multiple UEs can be achieved by 3GPP network-controlled replication mechanism thus to support traffic with stringent latency, reliability or redundancy requirements:

- Adoption of 5GS network to the industrial redundancy topology and forwarding packets with same MAC address to different redundant user plan paths
- Replication and elimination performed within 3GPP system for two different PDU Sessions, one per UE, assuming UE is independently registered
- Ability to trigger replication within the 3GPP system, protocol aspects to replicate/eliminate
- Enabling additional granularity (depending on traffic characteristics) for duplication of traffic over different redundant user plan paths
- Influence Multi-UE device for establishing one or two PDU Sessions for a certain application. In addition, provide rules to influence Multi-UE device for steering, switching and splitting the traffic

Project funded by Horizon Europe, Grant Agreement #101057083

- Extended capability exposure to allow Application Function to discover, request or influence selective duplication, 5GS reliability and redundancy functionality

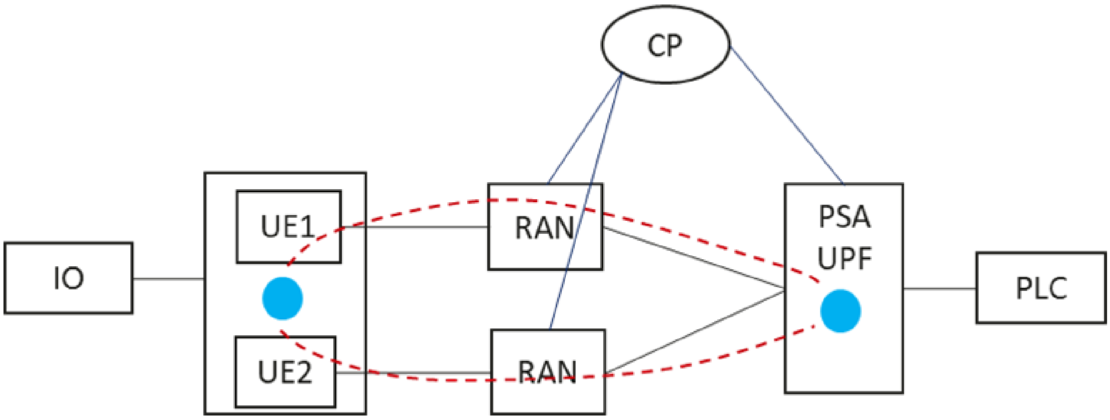

Figure 31 Support of native redundant UP paths without relying on upper layer protocols

3.3.5 Configuring a 5G network based on IEC 61499 industrial applications

IEC 61499 is an international programming standard for the design and implementation of PLC-based distributed control systems. Today, IEC 61499 quickly gains supporters in industry. One of the key advantages of IEC 61499 is that it supports the development of modular, distributed and flexible control systems. IEC 61499 programs are based on an object-oriented approach and consist of a graphical network of Function Blocks (FBs). Each type of FBs has an interface that includes input and output data and event ports. The inner logic of the FB is defined by Execution control Chat (ECC). ECC is an executable state machine. It is managed by input events and logic conditions.

The IEC 61499 standard supports a distributed software architecture for distributed industrial networked systems. Such systems consist of various actors, including actuators, smart sensors, programmable logic controllers, edge devices, and cloud-based services. IEC 61499 systems are based on networking principles; however, nowadays, the overall approach is that an IEC 61499 compatible device typically follows the parameters set by the network.

IEC 61499 devices are typically “passive”, as all communication parameters are provided by the network itself. However, in the case of 5G, the intricacies of the network configuration can have a significant impact on the characteristics of the distributed industrial control system, such as delays, accuracy, data logging, etc. This is just an introduction to the problem. We plan to explore the following tasks formulated below and present the related results in the final deliverable D3.5:

- Identify network configuration parameters that would be useful to configure directly from a IEC 61499 distributed applications;
- Develop an approach to configure industrial 5G networks from IEC 61499 distributed control software;
- Develop an interface of Service Interface Function Block (SIFB) for 5G network configuration.

3.3.6 Summary

As a summary of this section, some URLLC considerations for advance industrial communications have been proposed, in particular, padding compression to increase URLLC capacity, BAT adjustment and Survival Time enhancement in the UE-core network-the UE scenario, support of native redundant UP paths without relying on upper layer protocols. Besides, support integration of 5G Virtual Network and legacy LAN and exposure of 5G capability for factory areas have been considered. The proposed inputs can be considered as potential issues to be studied within the 3GPP standardization in Rel'19 and 20.

Besides, an introduction to configuration aspects of a 5G network based on IEC 61499 industrial applications was briefly mentioned in this section. However, this study is in progress and the main results of it will be presented in the final Deliverable D3.5.

4 OPC UA & 5G interworking: CAMARA Quality on Demand (QoD) interface for OPC UA integration

Different types of OPC UA sessions may require a differentiated treatment by the private 5G network. Providing applications with a tailored QoS connectivity service in a flexible manner is enabled in 5G by the Service Based Architecture (SBA) adopted in the design of the 5GCore (5GC). Following the SBA, each Network Function (NF) exposes its capabilities to the rest of NFs through well defined REST endpoints. In addition, the 5GC includes the Network Exposure Function (NEF) that is in charge of exposing a selected set of 5GC functionalities to other software entities outside the 5GC, for example to an application function (AF) that would like to request a specific quality of service from the 5GC.

QOS Profile labels	Network Service Description	Network Service Constraints
QOS_E	Latency stays stable under congestion (throughput up to a certain limit, e.g. 500kbps). The application bitrate should not exceed the limit.	The service is only provided, when the UE is located within the geographical area, according to the agreement with the CSP. The access to the network service is restricted to a maximal number of simultaneous session.
QOS_L	The 5G System throughput is prioritized up to certain higher limit (e.g. 20Mbps), or without an explicit limit. At high load, the throughput may be capped at the limit and can be degraded to lower effective throughput (without a minimum level).	The service is only provided, when the UE is located within the home network. The access to the network service is restricted to a maximal number of simultaneous session.
QOS_M	The 5G System throughput is prioritized up to certain medium limit (e.g. 8Mbps). At high load, the throughput may be capped at the limit and can be degraded to lower effective throughput (without a minimum level).	The service is only provided, when the UE is located within the home network. The access to the network service is restricted to a maximal number of simultaneous session.
QOS_S	The 5G System throughput is prioritized up to certain lower limit (e.g. 4Mbps). At high load, the throughput may be capped at the limit and can be degraded to lower effective throughput (without a minimum level).	The service is only provided, when the UE is located within the home network. The access to the network service is restricted to a maximal number of simultaneous session.

Figure 32. Defined CAMARA QoD QoS profile labels

This application triggered "Quality on Demand (QoD)" functionality, is not only relevant in the industrial domain, but also in other domains, such as multimedia, where QoD may be required for example by a video call that suffers from a temporary quality degradation. To address these type of use cases that already enjoy a high market demand, Mobile Network Operators (MNOs) recently joined forces around the CAMARA initiative [[31]], with the goal of defining common APIs that can be accessed by application developers. Hyperscalers, i.e. AWS, Azure and Google Cloud, are also part of the CAMARA initiative acting as aggregators that offer the CAMARA APIs to the application developers and connect to the appropriate MNO depending on the network used by the mobile device consuming those APIs. It is worth noting that CAMARA only defines the north-bound API exposed to the application developer, but does not standardize the backend required to implement said API. The APIs standardized by CAMARA at the time of writing include among others:

- Quality on Demand (QoD) [[32]]: Used by the AF to request from the network a specific QoS treatment.
- DeviceStatus [[33]]: Used by the AF to discover if the device is in connected or idle state, of it is roaming
- DeviceLocation [[34]]: Used by the AF to verify if the device is within a given geographical area.

The CAMARA QoD API is the most relevant one in terms of potential integration with OPC UA. The QoD API has defined a set of QoS profile levels that denote different quality levels, leaving to each MNO the implementation of the mechanisms required to enforce those quality levels. These QoS profile levels are described in Figure 32.

A potential technical gap to be pursued in the integration of the CAMARA QoD API in industrial scenarios is the definition of additional QoS profile labels that correctly capture the typical requirements of OPC UA connections. Each private network vendor would then have to provide a proprietary implementation of the mechanisms required to enforce the desired QoS level.

Figure 33 below depicts the full chain interactions involved in enforcing a given level of QoS for a specific OPC UA connection in the context of an industrial 5G network.

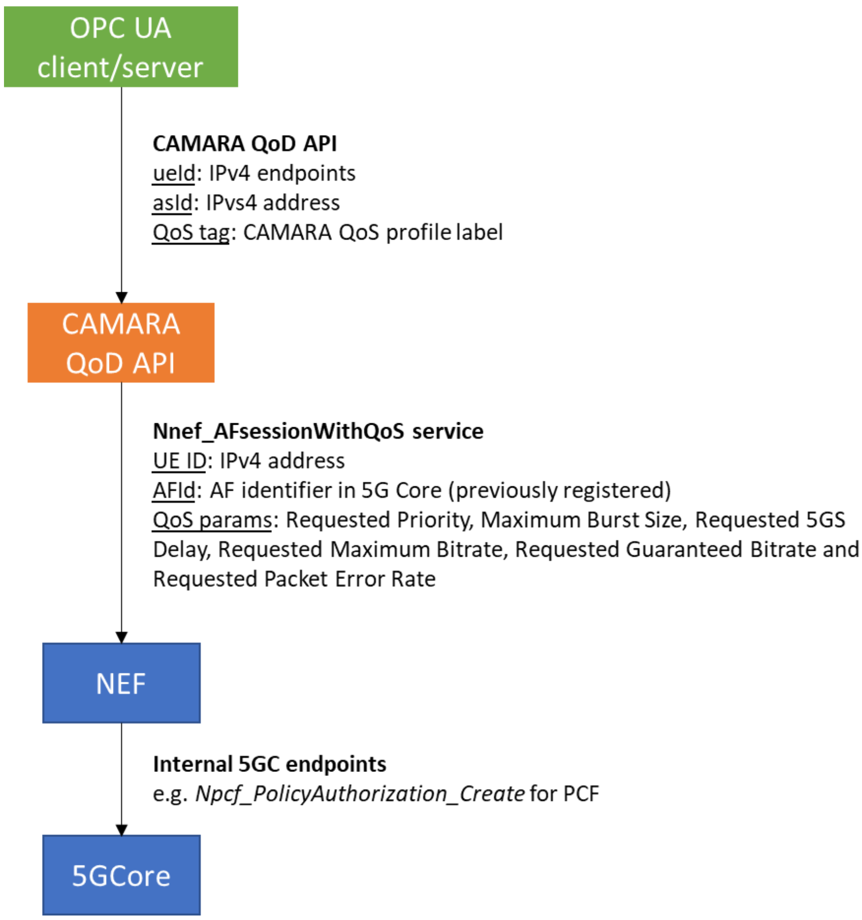

Figure 33. CAMARA QoD API in the context of OPC UA - 5GC integration

In particular, the procedure consists of the following steps:

- Step 1: The OPC UA client or server invokes the CAMARA QoD API using as identifiers the IPv4 address of the OPC UA client and the OPCUA server, and indicating the required QoS profile label, which needs to be known to the OPCUA devices a priori.
- Step 2: The CAMARA QoD API translates the high level QoD API call to a more detailed call to the *Nnef_AFSessionWithQoS_Service* exposed by the NEF of the private network. This call may require the CAMARA QoD API to resolve the IPv4 addresses to internal 3GPP identifiers (e.g. GPSI [[35]]), and will translate the CAMARA QoS profile label to actual QoS parameter values, such as priority, delay, bitrates, etc
- Step 3: Finally, the NEF will translate the previous call to an *Npcf_PolicyAuthorization_Create* request to the PCF, through which a specific Policy and Charging Control (PCC) rule will be created in the PCF, which will then be translated to specific configurations in the User Plane Function (UPF) and the gNB to enforce the required QoS.

5 Conclusion

This section summarizes the main observations, gaps, novelties, innovations defined in this Deliverable in comparison with SoA related to integration of 5G network to industrial communications as follows:

1. Existing 5G-ACIA Standalone and Public Network Integrated Non-Public Networks (PNI-NPN) deployment options where analysed together with their advantages and drawbacks. As a result, it was found that there is not a PNI-NPN option that can support sensitive traffic from industrial applications locally within OT domain without traversing a third party network (i.e. core network of Mobile Network Operator) or without installation of User Plane Function equipment in OT domain under control of the third party, i.e. MNO.

   To address this gap, Deliverable a **new PNI-NPN deployment option** was proposed in the Deliverable. Such solutio provides OT sensitive traffic localisation on-premise by means of special programmable traffic steering switch that has 3GPP tunnel encapsulation and decapsulations functions. Traffic rules for localizing of sensitive data go from edge orchestrator belonging to the OT. That is, all sensitive traffic is under control the OT, not an MNO. In addition, this new deployment option has a short data traffic transmission path in comparison with other 5G-ACIA PNI-NPN options that makes it applicable to support industrial use cases with delay critical requirements.

2. The three main existing industrial Ethernet protocols, namely, EtherCAT, Ethernet/IP, and PROFINET that are used in the layer-2 infrastructure of OT production domain and IT enterprise domain were analysed from the viewpoint of their technical features and challenges related to their integrations with the 5G system. As a result of this analysis, **the 5G enhancements for integration with EtherCAT, Ethernet/IP, and PROFINET** as the basic industrial protocols were formed.

3. For existing 3GPP 5G-TSN integration model and related state of the art of scheduling techniques **a novel Configured Grant (CG) scheduling scheme** was proposed. The designed scheduling scheme coordinates its scheduling decision with the global scheduling performed in TSN to meet efficiently and effectively the end-to-end latency and determinism requirements of TSN traffic. In particular, the proposed scheduling scheme exploits the information provided by TSN about the characteristics of the TSN traffic (periodicity, packet size, etc.) to satisfy the maximum latency established by the TSN scheduling decision.

   The proposed scheme allows avoiding scheduling conflicts that can arise with common CG scheduling schemes that allocate radio resources periodically for each TSN flow based on the periodicity of the packets that may result in the allocation of the same radio resources for the transmission of packets of different TSN flows. The proposed CG scheduling scheme solves this issue by configuring several grants for each TSN flow in order to avoid radio resource allocation conflicts among different TSN flows.

The achieved performance evaluation results validate that the proposed scheduling scheme considerably increases the number of TSN flows that can be satisfactorily served in the integrated 5G-TSN network compared to common CG schemes.

4. Some observations were considered to resolve some existing 3GPP challenges for industrial communications.

   In particular, **the enhancements to improve efficiency of transmission capacity and the URLLC performance** (lower latency and higher reliability) in the industry communications should include:

   - Support Burst Arrival Time (BAT) adjustment and survival time in device-network-device scenario;
   - Support redundant user plane paths based on multiple UEs without relying on application layer duplication redundancy protocols;
   - Support to avoid transferring unnecessary bits (e.g. paddings) in 5GS to improve the efficiency of transmission capacity and the latency performance;
   - Improve the support of redundant user plane paths with adaptive mechanism for optimal usage of 5GS resources and to avoid the duplicated reliability.

   Besides, **the enhancements to support integration of 5G VN and legacy LAN** should include:

   - Integration architecture, e.g. cross-site tunnel establishment and release, QoS provisioning and monitoring for cross-site tunnel, intelligent traffic steering on cross-site tunnel, exposure for reporting and configuration mapping;
   - Management of the DN side devices and the UE side device as 5G VN group members;
   - Support loop-free Ethernet topology in the case of Ethernet PDU Sessions;
   - Support NEF exposure of 5G capabilities to fulfill industrial applications' requests: e.g. Acknowledging a AF service request within indicated time, VLAN tag assignment without AAA-Server, UE-to-UE QoS with one NEF API, Per QoS characteristic monitoring, IP address assignment via NEF exposure.

     All these enhancements are supposed to push within WP7 towards the 3GPP standardization studies in Rel'19 and 20.

5. For OPC UA and 5G interworking it was proposed to use **CAMARA Quality on Demand interface with defined a set of QoS profile levels**. In particular, it was developed a procedure to enforce a given level of QoS for a specific OPC-UA connection in the context of an industrial 5G network.

This is an intermediate deliverable and as future plans we are going to extend the results presented above and consider the following aspects to include in the final deliverable, D3.5:

– 5G user plane connectivity scenarios for wired and wireless devices via legacy L2/L3 OT infrastructure for existing 5G-ACIA deployment options and new option proposed to this deliverable;

– Enhancements for the scheduling scheme proposed in this deliverable to ensure that it can be properly implemented in current 5G networks. In addition, it is supposed to analyze how much current solution could be improved using optimization techniques;

- Different issues of 5G and OT domain integration, namely, further integration of industrial Ethernet protocols with 5G networks based on observations obtained in this deliverable;
- OPC-UA and 5G integration to explore synergies with the GSMA Open Gateway initiative and the CAMARA framework started in this deliverable;
- Configuration of a 5G network based on IEC 61499 industrial applications that was initially introduced in this deliverable, in particular, identification of network configuration parameters from IEC 61499 distributed applications, develop an approach to configure these parameters and an interface based on Service Interface Function Block (SIFB) for 5G network configuration;
- Other aspects related to enhancing 5G non-private network communications.

References

[1] 5G-ACIA White Paper, “5G Non-Public Networks for Industrial Scenarios”, published in July 2019.

[2] ETSI white paper, WP#28, “MEC in 5G networks”, June 2018.

[3] 5G-ACIA White Paper “Integration of Industrial Ethernet Networks with 5G Networks”, Nov. 2019.

[4] 5G-ACIA White Paper “Industrial 5G Edge Computing –Use Cases, Architecture and Deployment”, Feb. 2023.

[5] Netmanias technical blog posted by Harrison J. Son, “7 deployment scenarios of private 5G networks”, Oct 2019.

[6] F. Paolucci et al, “User Plane Function Offloading in P4 switches for enhanced 5G Mobile Edge Computing”, in DRCN proceedings, Italy, 2021.

[7] S. S. Craciunas et al, “Scheduling real-time communication in IEEE 802.1Qbv time sensitive networks”, ACM Int. Conf. Proceeding, 2020.

[8] IEEE, IEEE Std 802.1Qcc: Standard for Local and Metropolitan Area Networks--Bridges and Bridged Networks, 2018.

[9] 3GPP, System Architecture for the 5G System (5GS); Rel’ 18, TS 23.501, V18.2.2, 2023.

[10] IEEE, IEEE 802.1Q: Standard for Local and metropolitan area networks -- Bridges and Bridged Networks, 2018.

[11] Z. Satka et al, “QoS-MAN : A Novel QoS Mapping Algorithm for TSN-5G Flows”, the 28th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA'22), Taiwan, 2022.

[12] Y. Zhang et al, “QoS-Aware Mapping and Scheduling for Virtual Network Functions in Industrial 5G-TSN Network,” 2021 IEEE Glob. Commun. Conf. GLOBECOM, 2021.

[13] M. J. Yang et al, “An Uplink Transmission Scheme for TSN Service in 5G Industrial IoT,” Int. Conf. ICT Converg., 2020.

[14] 3GPP, TSG Services and System Aspects; Service requirements for the 5G system; Rel’ 19, TS 22.261, V. 19.3.0, 2023.

[15] D. Ginthör et al, “5G RAN slicing for deterministic traffic”, in IEEE Wireless Communications and Networking Conference (WCNC), China, 2021.

[16] R. B. Abreu et al, “Scheduling Enhancements and Performance Evaluation of Downlink 5G Time-Sensitive Communications,” IEEE Access, 2020.

[17] Y. Zhang et al, “Wireless/wired integrated transmission for industrial cyber-physical systems: risk-sensitive co-design of 5G and TSN protocols,” Sci. China Inf. Sci., 2022.

[18] Nokia, “5G plug-and-produce. How the 3GPP 5G System facilitates Industrial Ethernet deployments”, white paper, 2020.

[19] A. Larrañaga Zumeta, M. C. Lucas-Estañ, S. Lagén, Z. Ali, I. Martinez, and J. Gozalvez, “An open-source implementation and validation of 5G NR configured grant for URLLC in NS-3 5G LENA: A Scheduling case study in Industry 4.0 scenarios,” SSRN Electron. J., Aug. 2022.

[20] K. Montgomery et al, “Wireless User Requirements for the Factory Workcell", NIST Advanced Manufacturing Series 300-8, Nov. 2020.

[21] Z. Satka et al., “Developing a Translation Technique for Converged TSN-5G Communication,” IEEE Int. Work. Fact. Commun. Syst. - Proceedings, WFCS, 2022.

[22] D. Pannell, “Choosing the Right TSN Tools to meet a Bounded Latency”, in IEEE SA Ethernet & IP @ Automotive Technology Day, Detroit, September, 2019.
[23] “MA734 | 8-Bit to 12.5-Bit, 3µs Low-Latency Contactless Angle Sensor | MPS.” [Online]. Available: https://www.monolithicpower.com/en/ma734.html [Accessed: 13-Nov-2022].
[24] R. Automation, “1756 ControlLogix I/O Modules Specifications Technical Data”, available at https://literature.rockwellautomation.com/idc/groups/literature/documents/td/1756-td002_-en-e.pdf
[25] R. Automation, “Estimated Logix 5000 Controller Instruction Execution Times Reference Manual”, available at https://literature.rockwellautomation.com/idc/groups/literature/documents/rm/logix-rm002_-en-p.pdf
[26] 3GPP, TSG RAN; Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC); Rel’ 16, TR 38.824. V16.0.0, 2019.
[27] 3GPP, TSG RAN; Base Station (BS) radio transmission and reception; Rel’ 17, TS 38.104, 2021.
[28] 3GPP, TSG RAN; Physical layer procedures for data; Rel’ 16, TS 38.214, V16.1.0, 2020.
[29] 3GPP, TSG RAN; Medium Access Control (MAC) protocol specification; Rel’ 16, TS 38.321, V16.6.0, 2021
[30] 5G-ACIA white paper, “Exposure of 5G Capabilities for Connected Industries and Automation Applications”, Version 2, March 2021.
[31] CAMARA project. Available at: https://github.com/camaraproject
[32] https://github.com/camaraproject/QualityOnDemand
[33] https://github.com/camaraproject/DeviceStatus
[34] https://github.com/camaraproject/DeviceLocation
[35] https://github.com/camaraproject/WorkingGroups/blob/main/Commonalities/documentation/UE-Identification.md