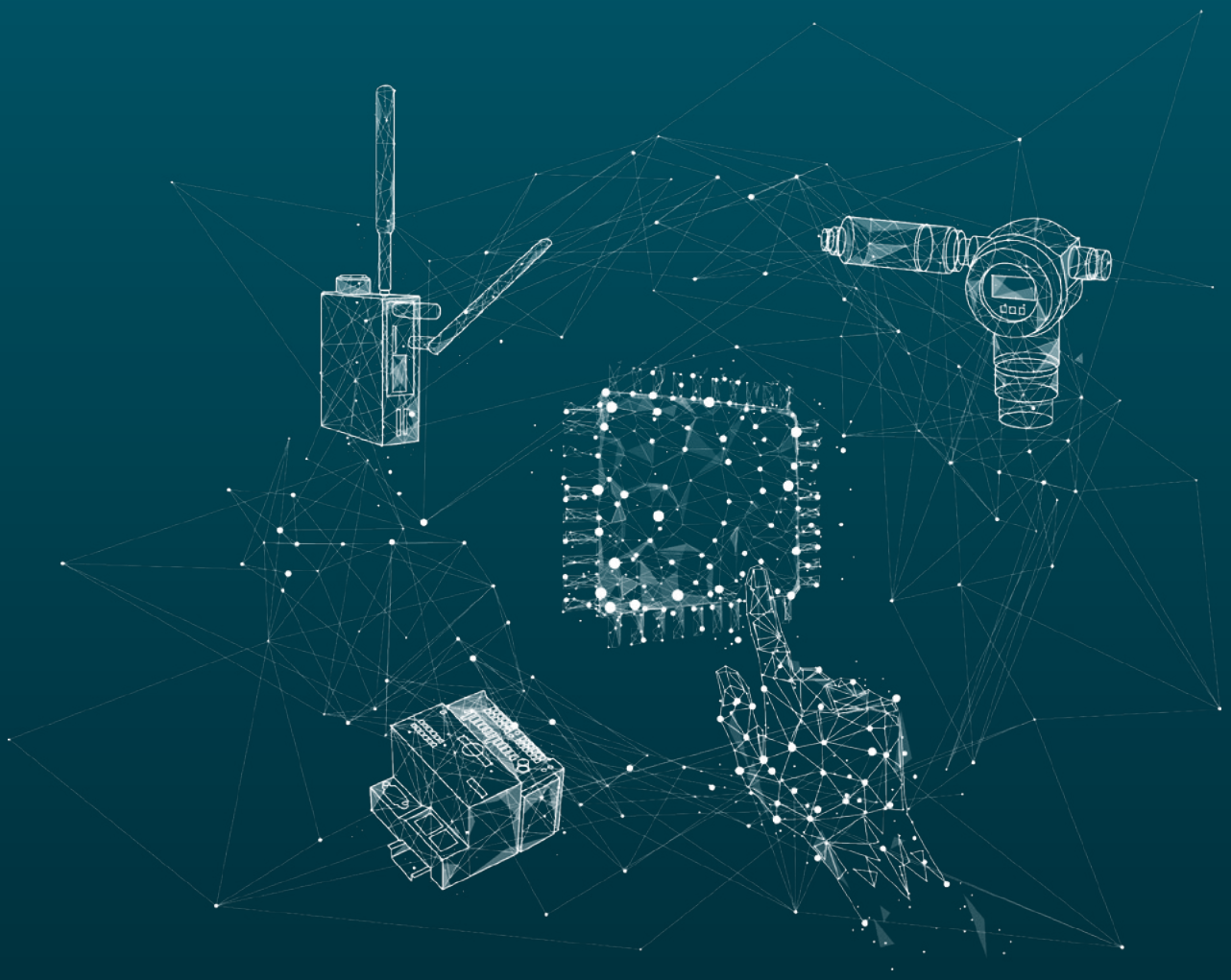
5G-ACIA White Paper

Industrial 5G Devices – Architecture and Capabilities



5G Alliance for Connected Industries and Automation

Table of Contents

1. Executive Summary 5

2. Introduction 6

3. Industrial 5G Devices 7
- 3.1 Types of Industrial 5G Devices 7
  - 3.1.1 Low-Latency Sensors/Actuators 8
  - 3.1.2 Low-Power Sensors/Actuators 8
  - 3.1.3 2D/3D Sensors 8
  - 3.1.4 HMI and xR 8
  - 3.1.5 PLCs and Controllers 9
  - 3.1.6 Gateways 9
  - 3.1.7 TSN Ports 10
- 3.2 Characteristics of Industrial 5G Devices 11
  - 3.2.1 Time Characteristics 11
  - 3.2.2 Data Characteristics 12
  - 3.2.3 Power Characteristics 13
  - 3.2.4 Time Synchronization 13
  - 3.2.5 Positioning 13
  - 3.2.6 Communication Themes 14
- 3.3 Examples of Industrial 5G Devices 14
  - 3.3.1 5G IP67 Sensor 16
  - 3.3.2 5G Smart Sensor 16
  - 3.3.3 5G IIoT Level Sensor 18
  - 3.3.4 5G Dual-Channel Adapter 18
  - 3.3.5 5G Remote I/O for Process Control 19
  - 3.3.6 5G Process Control via Mobile Panel 19
  - 3.3.7 Mobile App for 5G Industrial Devices for Augmented Field Applications 20
  - 3.3.8 5G Drone Operation 20
  - 3.3.9 5G Ethernet Bridge 21
  - 3.3.10 5G Wireless Router 21
  - 3.3.11 5G Industrial Gateway 22
  - 3.3.12 5G Mobile Tracker 22
  - 3.3.13 5G Valve Terminal 23
  - 3.3.14 5G Controller (Remote I/O) 24

4. Logical Reference Architecture for Industrial 5G Devices 25
- 4.1 Top-Level Logical Architecture 25
- 4.2 Practical Logical Architecture 26
  - 4.2.1 Logical Architecture for Supporting Applications Inside a 5G Industrial Device 27
  - 4.2.2 Logical Architecture for Supporting Applications or Networking Using IP or Ethernet with Traditional Non-Time-Aware QoS 28
  - 4.2.3 Logical Architecture for Supporting Applications Using IP and Ethernet with QoS and Precision Time Protocol over a 5G Radio Link 29
  - 4.2.4 Logical Architecture for Supporting Applications Using Ethernet with IEEE TSN 31

4.3 Device Authentication 32
4.3.1 Introduction 32
4.3.2 Primary Authentication for PNI-NPNs 32
4.3.3 Primary Authentication of SNPNs 32
4.3.4 NSSAA and Secondary Authentication 34
4.3.5 Summary 34

5. Industrial 5G Device Physical Reference Architecture 35
5.1 Explosion Protection for Devices in Hazardous Areas 35
5.1.1 Introduction 35
5.1.2 Classification of Zones 35
5.1.3 Types of Explosion Protection for Industrial Devices 35
5.2 Physical Implementation for Storing Credentials 37
5.2.1 Removable Secure Element 38
5.2.2 Embedded Secure Element Without Key Management Interface 38
5.2.3 Embedded Secure Element with Key Management Interface 38
5.2.4 Provisioning of Cellular Credentials 38
5.3 Chipset Versus Module 39
5.4 Radio Module Form Factor Standards 39
5.5 Standalone Versus Integrated Application Processor 39
5.6 Interface Between Application Processor and Radio Module 40
5.6.1 Data Interface 40
5.6.2 Time Synchronization Interface 40
5.7 Generic Block Diagrams for Industrial 5G Devices and Interface Options 40
5.7.1 Low-Power and Low-Latency Sensors/Actuators, 2D- and 3D-Sensor Industrial 5G Devices 41
5.7.2 HMI and xR Devices 41
5.7.3 Gateways and PLCs/Controllers 42
5.7.4 TSN Port Industrial 5G Devices 43

6. Conclusions 45

7. Definitions of Acronyms and Key Terms 46

8. References 49

1. Executive Summary

This white paper provides an overview of the kinds of devices that can be needed in order for 5G to benefit the manufacturing industry and related sectors. As 5G systems are implemented in factories and other settings, attention is increasingly shifting to designing devices that will let them work on the shop floor. A whole new generation of 5G-compatible devices is now being developed. This paper provides an introduction and practical guide to this field for everyone who is directly or indirectly involved in it, whether they are academics, manufacturers, factory owners or operators, designers, or engineers. Its main purpose is to provide an easy-to-read overview of the various categories of devices and solutions that are now appearing, while going into greater technical detail on key technical topics and design issues.

The main types of 5G devices are presented and described and a number of real-world examples discussed while describing the most important technical issues, challenges, and solutions involved in each case. On a more theoretical level, reference architectures are then presented for the most common types of industrial 5G devices, including generic block diagrams.

Finally, various aspects of the physical architecture of such devices are discussed, covering challenges such as explosion protection, storage of credentials, the pros and cons of chipset versus module solutions, radio module form factor standards, a comparison of standalone and integrated application processors, and implementation of interfaces.

About 5G-ACIA

The **5G Alliance for Connected Industries and Automation** (5G-ACIA) was established to serve as the main global forum for addressing, discussing, and evaluating relevant technical, regulatory, and business aspects of 5G for the industrial domain. It embraces the entire ecosystem and all relevant stakeholders, which include but aren't limited to the operational technology industry (industrial automation companies, engineering companies, production system manufacturers, end users, etc.), the information and communication technology industry (chip manufacturers, network infrastructure vendors, mobile network operators, etc.), universities, government agencies, research facilities, and industry associations. 5G-ACIA's overarching goal is to promote the best possible use of industrial 5G while maximizing the usefulness of 5G technology and 5G networks in the industrial domain. This includes ensuring that ongoing 5G standardization and regulatory activities adequately consider relevant interests and requirements and that new developments in 5G are effectively communicated to and understood by manufacturers.

2. Introduction

The fifth-generation standard for broadband cellular networks (5G) enables reliable, low-latency, high-bandwidth data transmission, making it a key technology for the future of industrial communications. The introduction of 5G to factories and a wide range of other industrial facilities is also creating a need for industrial devices that support the 5G standard.

How should an industrial 5G device be designed? This white paper provides chip manufacturers, module vendors, and device manufacturers with guidance on the available choices.

Chapter 3 describes various kinds of industrial 5G devices, mainly from an operational technology (OT) perspective. It also contains a large collection of example industrial 5G devices gathered from 5G-ACIA members. Chapters 4 and 5 describe the logical and physical architecture of industrial 5G devices.

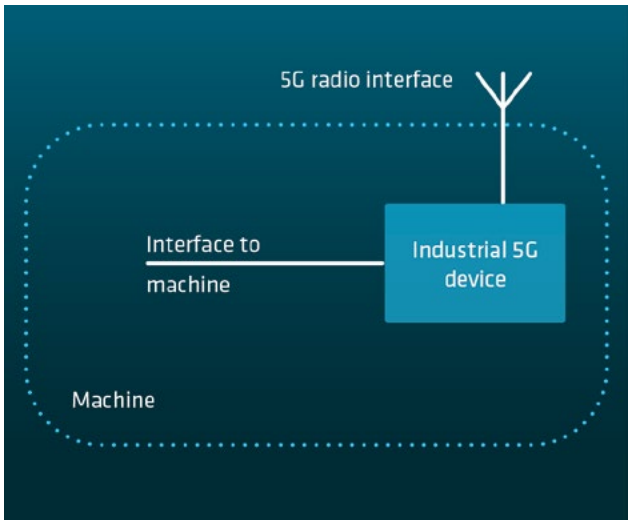
This white paper makes it clear that the field of industrial 5G devices draws on a wide range of engineering disciplines including operational technology (OT) and information and communication technology (ICT). It also integrates aspects of mechanical design, product safety, and cybersecurity.

3. Industrial 5G Devices

Industrial 5G devices come in a wide variety of types and shapes and can be deployed for diverse use cases as described in [4] and [8]. Section 3.1 provides an overview of different industrial 5G device types, section 3.2 describes some of their characteristics, and section 3.3 presents various example applications.

The following discussion includes references and links to example use cases and related requirements. For the sake of conciseness, it only goes into detail on a relatively small number of use cases for applications that include motion control, portable tools in assembly areas, remote augmented reality, and process automation. The numerical values and ranges given in section 3.2 for industrial devices in certain use cases are only examples.

Figure 1: An industrial 5G device as part of a machine



Industrial 5G devices can be either standalone or integrated into something else. Figure 1 shows an industrial 5G device integrated in a machine. This approach makes it possible to depict an industrial 5G device while showing only the functions that are most relevant from a communication perspective.

3.1 Types of Industrial 5G Devices

This section presents seven different types of industrial 5G devices:

- Low-latency sensors/actuators
- Low-power sensors/actuators
- 2D/3D sensors
- HMI and xR
- PLCs and controllers
- Gateways
- TSN ports

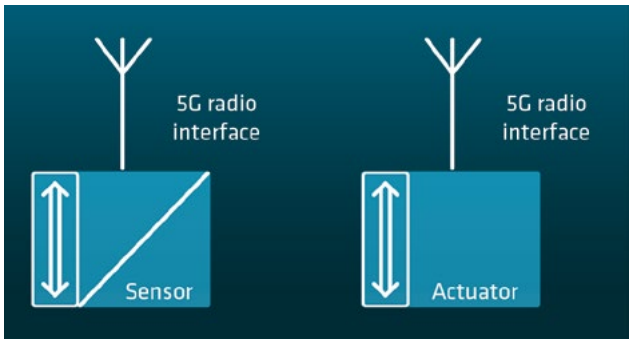
These industrial 5G devices are described from an operational technology perspective. Their types are indicated when discussing their logical and physical architectures.

The gateway industrial 5G device discussed in 3.1.6 involves transparent information transfer between different communication technologies on various protocol levels. It integrates industrial protocol gateway, IP routing, and Ethernet bridging functionality. The TSN port industrial 5G device (forming part of a distributed TSN bridge within the 5G system) is separately described in 3.1.7 since it has a different architecture.

The industrial 5G device types discussed here are illustrated by a large collection of examples in section 3.3.

3.1.1 Low-Latency Sensors/Actuators

Figure 2: Low-latency sensor and actuator

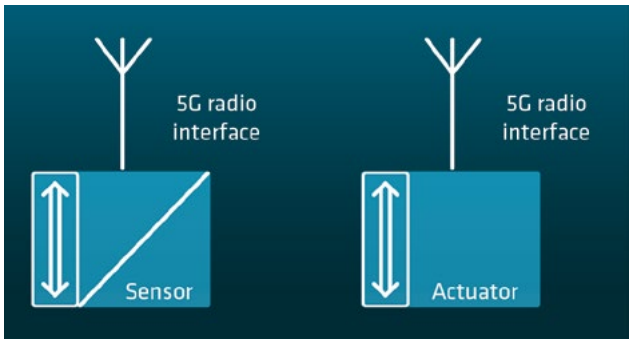


Low-latency sensors and actuators are normally wired, but in 5G they can also be connected via a radio interface to a PLC and/or controller in the cellular network. In this case, real-time communication and high reliability are essential.

These devices are commonly deployed in mobile robot use cases, many of which involve low-latency communication. This statement also applies to interactions with stationary peripherals and cooperation with other robots.

3.1.2 Low-Power Sensors/Actuators

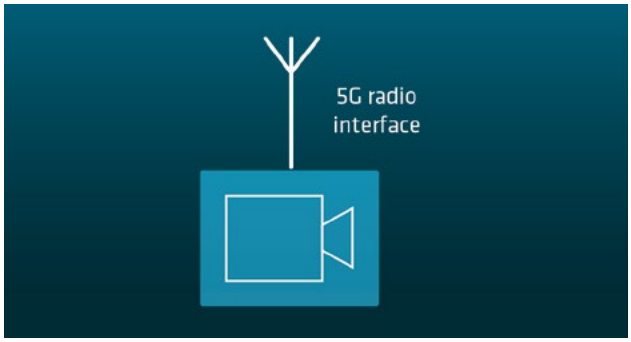
Figure 3: Low-power sensor and actuator



A low-power sensor or actuator has a radio interface to the cellular network. These devices are typically used for monitoring condition, productivity, or production quality. They can be battery-powered and may spend much of the time in sleep mode. Since they are typically expected to operate for several years without recharging, it's essential for them to be energy-efficient.

3.1.3 2D/3D Sensors

Figure 4: 2D/3D sensor



2D/3D sensors capture two- and/or three-dimensional data from an industrial manufacturing facility or process. They have a radio interface to the cellular network and can include cameras and LIDARS, for example, and deliver 2D and/or 3D images at defined frame rates.

2D/3D sensors are typically used to collect production data that are then analyzed by an AI-based system. One application is data collection for quality assurance and another is fine-grained positioning.

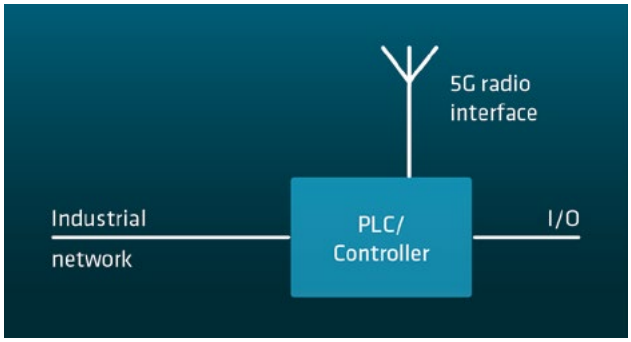
3.1.4 HMI and xR

Figure 5: HMI and xR

In the context of industrial 5G, HMI or extended reality (xR) can be used to provide a user interface to a manufacturing facility or process. This involves a radio interface to the cellular network as well as communication media that can include video screens, loudspeakers, cameras, and/or microphones. Their purpose is typically to provide visual information to an operator for interacting with an industrial facility or process.

3.1.5 PLCs and Controllers

Figure 6: PLC



A PLC/controller (PLC stands for "programmable logic controller") has a radio interface to the cellular network, another interface to one or more local industrial networks, and/or various I/O interfaces. It is basically an industrial computer that is used to control one or more processes.

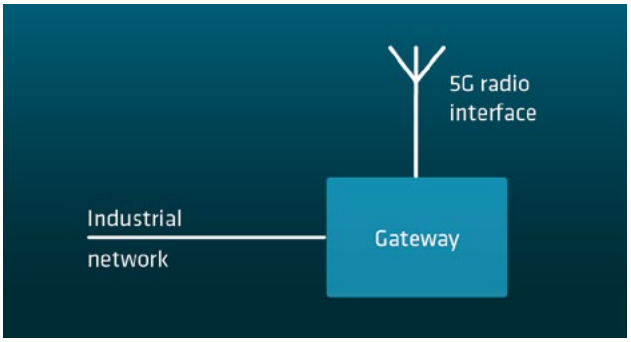
A 5G radio interface is typically connected to one or more of the following:

- A supervisory system
- Another PLC or other controller
- Devices in the control loop

When a 5G radio interface is used to communicate with devices in the control loop, another PLC, or some other type of controller, communication is time-critical. Outside of control loops, the timing requirements are less strict.

3.1.6 Gateways

Figure 7: Gateway



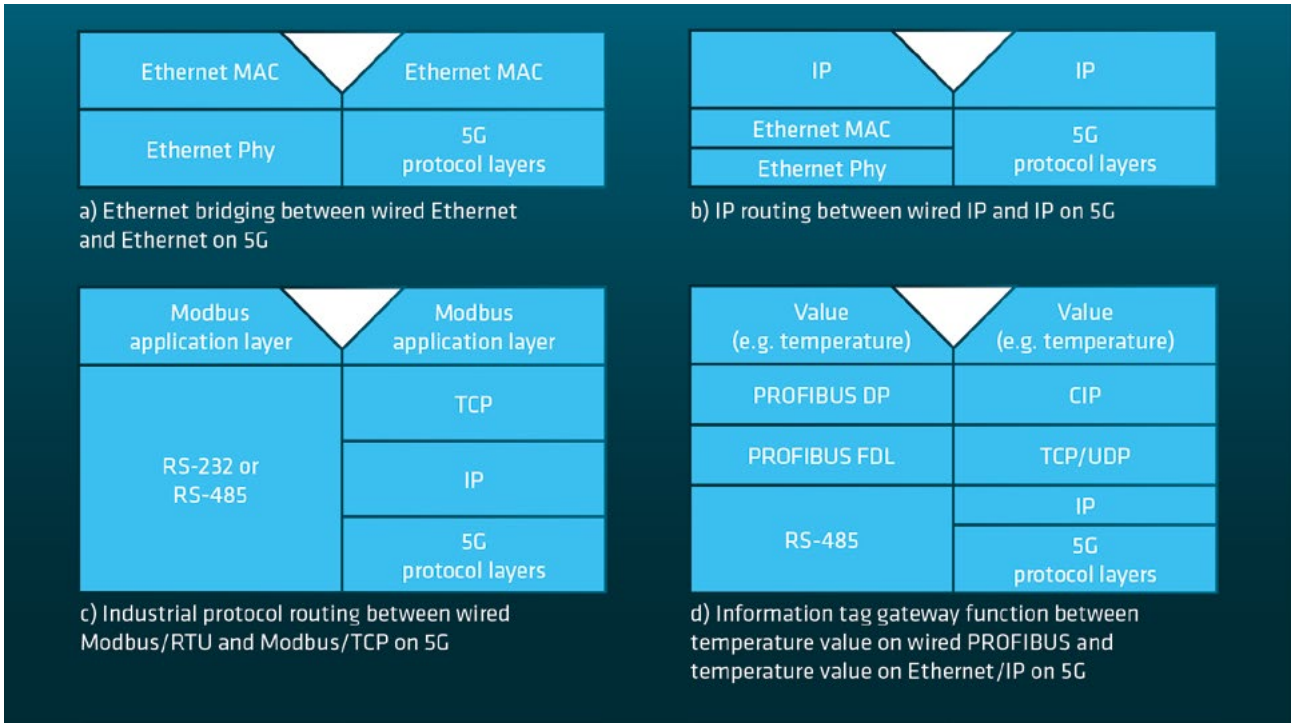
A gateway has a radio interface to a cellular network and a standardized wired (or wireless) interface to an industrial network. Its purpose is to relay information between the two.

Common industrial network interfaces include industrial Ethernet and fieldbus interfaces.

A gateway can operate in different protocol layers; figure 8 shows some examples.

In the context of industrial 5G, HMI or extended reality (xR) can be used to provide a user interface to a manufacturing facility or process. This involves a radio interface to the cellular network as well as communication media, which can include video screens, loudspeakers, cameras, and/or microphones. Their purpose is typically to provide visual information to an operator for interacting with an industrial facility or process.

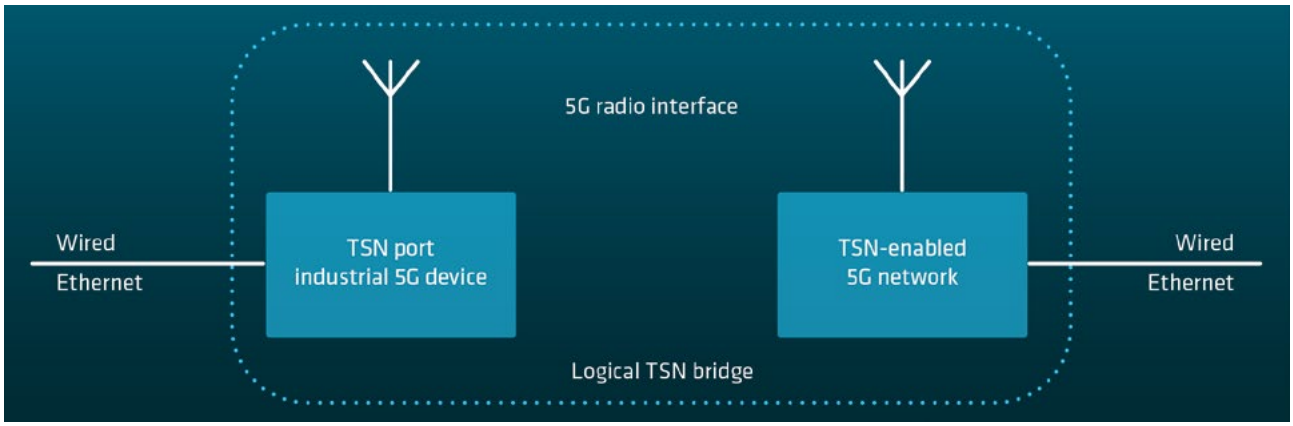
Figure 8: Examples of gateway functionality in different protocol layers



A 5G gateway can be preferrable to sensors and actuators with integrated 5G when retrofitting, in certain kinds of installations such as production/ process modules, and in challenging environmental conditions such as hazardous areas.

3.1.7 TSN Ports

Figure 9: A TSN port as part of a logical TSN bridge



An industrial 5G device can serve as a port in a distributed 5GS Ethernet bridge anchored to a 5G user plane function (UPF). A 5GS Ethernet bridge can be configured to support features and management interfaces that comply with the IEEE time-sensitive networking (TSN) standards and the generalized precision time protocol (gPTP, IEEE 802.1AS) for integration in TSN- and gPTP-capable Ethernet networks. This is explained in greater detail in chapter 4.

These devices can be employed, for instance, in mobile robots that need to interact with one another, collaborative robots (cobots) that grasp and hand over parts, and cooperative driving scenarios. In all of these cases, it's essential to synchronize the actions of multiple actors.

3.2 Characteristics of Industrial 5G Devices

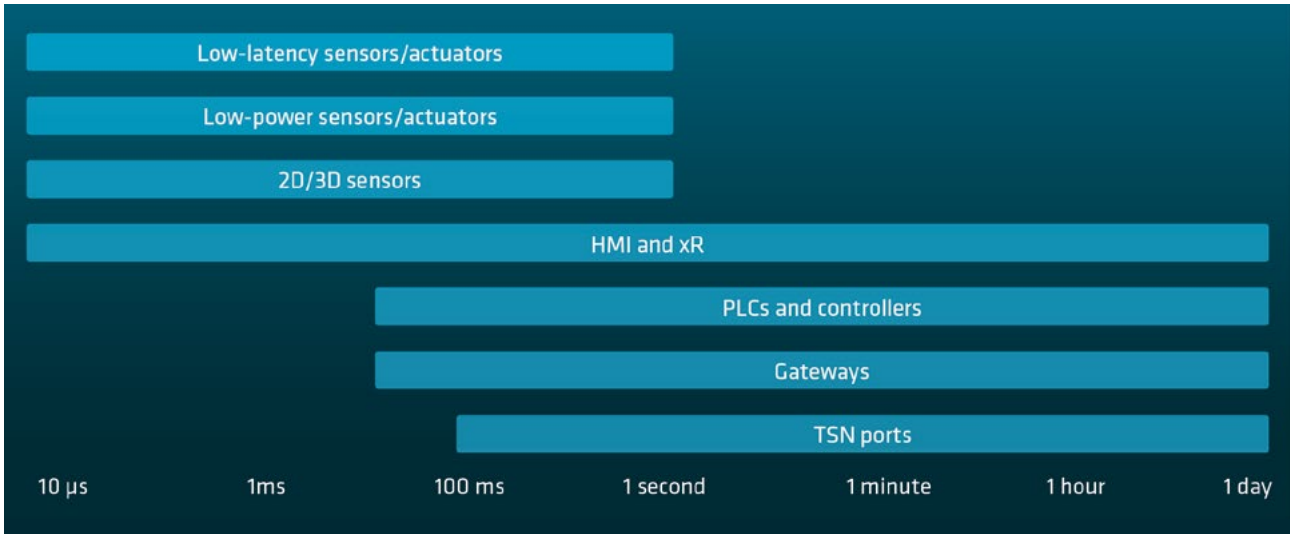
3.2.1 Time Characteristics

The factory automation protocols used for communication between a PLC and multiple devices follow a deterministic cyclic (or periodic) transmission pattern in which all of the sensors are read and all of the actuators are set during each cycle.

3GPP 5G uses the terms "transfer interval" and "periodic deterministic communication" to describe these patterns. The transfer interval is the time difference between two consecutive transfers of application data from an application to a 3GPP system via a service interface [8].

Figure 10 illustrates typical transfer intervals. Periodic deterministic communication (cyclical traffic) predominates in PLC/controller, low-latency sensor/actuator, and TSN port industrial 5G devices. There can also be aperiodic traffic such as alarms and firmware upgrades, which aren't included in the figure.

Figure 10: Examples of gateway functionality in different protocol layers



In addition to conventional devices deployed for factory automation, other devices are used to support Industrial IoT and Industry 4.0 in industrial installations. They include low-power sensors and actuators, 2D and 3D sensors, and HMI and xR devices. These devices typically exchange information at regular time intervals much longer than those for factory automation devices.

Gateways can be used with both conventional factory automation devices and the devices mentioned in the previous paragraph.

The latency requirements are largely determined by the cycle times and transfer intervals of the relevant factory automation protocols and use cases. The maximum permissible latency must be shorter than the transfer interval [8]. In isochronous use cases, the network latency may not exceed 20% to 50% of the cycle time or transfer interval [2].

The transfer time interval depends on the use case. To illustrate this, the interval for a mobile robot moving between two points depends on its navigation mode [7]:

- For infrastructure, track-guided navigation involves a transfer time of around 500 ms.
- Sensor/camera-based navigation involves a transfer time in the range of 10 to 100 ms.
- Cooperative driving requires a very short transfer time of around 5 ms.

The relationship between the transfer time interval and the required maximum network latency is different for low-power sensors and actuators, 2D and 3D sensors, and HMI and xR devices. For example, a 4k camera with a frame rate of 60 frames per second delivers data every 17 ms, but it is often acceptable for the network to have a greater latency than this.

3.2.2 Data Characteristics

Table 1 shows typical message sizes for various industrial 5G devices.

The smallest sensor or actuator data unit is a single bit, the value of which can indicate an input or an output. Analog values are commonly expressed as 16- or 32-bit values. Many sensors and actuators output multiple values, however, and protocol data are also communicated. The minimum frame size in Ethernet is 64 bytes, which is also the minimum message size as shown in table 1. The maximum message size is assumed to be 1522 bytes, corresponding to the largest Ethernet frame size with VLAN tagging. The required data rates can be calculated from the transfer intervals given in section 3.2.1.

Table 1: Typical data parameters of industrial 5G devices

	Message size	**Streams**	**Bitrate**
Low-latency sensors/ actuators	64 to 1522 bytes	1	≤ 200 kbit/s to 2 Mbit/s
Low-power sensors/ actuators	64 bytes or more	1	A few kbit/s to 2 Mbit/s
PLCs and controllers	64 to 1522 bytes	≥ 1	Up to line speed (100 Mbit/s, 1 Gbit/s)
Gateways	64 to 1522 bytes	≥ 1	Up to line speed (100 Mbit/s, 1 Gbit/s)
TSN ports	64 to 1522 bytes	≥ 1	Up to line speed (100 Mbit/s, 1 Gbit/s)

The bitrates given in table 1 correspond to the transmission speeds of active industrial 5G devices.

The data characteristics for PLCs and other controllers, TSN ports, and gateway industrial devices depend on the underlying use cases. With mobile robots, for example, different aspects can play a role depending on the functionality involved [7].

When the robots are moving between two points, the traffic models differ depending on the type of navigation used:

- Infrastructure- or track-guided navigation: a packet size of around 250 bytes and a data rate of 50 to 250 kbit/s
- Sensor- or camera-based navigation: a packet size of around 1500 byte and a data rate of 60 Mbit/s
- Cooperative driving: a packet size of around 250 bytes and a data rate of 125 kbit/s
- Interactions with stationary peripherals (grasping of unsorted piles) and a burst of 50 messages: a packet size of 1500 bytes and a data rate of around 400 Mbit/s

The data volumes generated by a 2D sensor depend on its resolution, the frame rate, the color depth, and any applied compression. For example, a 4k video with 60 frames per second and 24 bits per pixel has an uncompressed bitrate of 11.9 Gbps. A video stream can be compressed using a generic or application-specific algorithm.

Say that a 4k video camera is used to monitor product quality in a production process. Instead of sending all of the video frames to a central server, an application-specific algorithm can be used to select only those frames that actually show each new product captured. This can dramatically reduce the data stream.

3D sensors generate even more data than 2D sensors. Both 2D and 3D sensor data can be compressed using either generic or application-specific algorithms. Both types are generally also transmitted in the uplink direction.

The traffic characteristics of HMI and xR devices vary greatly depending on the use case. At the high end, video is streamed to a device at a bitrate that is generally between one and 25 Mbit/s. HMI and xR devices mainly transmit data in the downlink direction.

In automated processing plants, traffic is deterministic and periodic. Section 3.3.5 presents an example of remote I/O for process control.

3.2.3 Power Characteristics

One of the main reasons to deploy private industrial 5G networks is to make factories more flexible. More of the machines and devices used become wireless and battery-operated as a result.

HMI and xR devices are normally battery-operated. A typical use case is when a worker uses one or more devices throughout a shift. At the end of the shift, they are placed in chargers. This presupposes that the battery of each HMI or xR device has sufficient capacity to operate during an entire shift, which typically lasts about 10 hours including breaks. The same considerations apply to portable tools.

Low-power sensors and actuators are also usually battery-operated. The main reason for taking this approach is to reduce the cost of wiring. Batteries can be either rechargeable or disposable. In some cases, an entire device is discarded along with its battery.

Most other industrial 5G devices are typically powered by the machine they are installed on. For example, a gateway could be mounted on an AGV. In this case, the gateway is powered by the AGV's battery. The same considerations apply to mobile robots.

3.2.4 Time Synchronization

All industrial 5G devices need to be synchronized with different time domains. These include working clock domains and global clock domains. There is also a 5G clock domain, which is needed for 5G radio communication.

A working clock domain is needed for synchronizing sensors and actuators that are part of a control loop. Examples are robot collaboration and cooperative driving, in which time synchronization is paramount. Time synchronization can be explicit using protocols such as PTP (IEEE 1588) or gPTP (IEEE 802.1AS), or else implicit with read and write commands received from the PLC.

A global time domain is needed for sequences of events, time stamping of data, and time stamping of diagnostic events. It is usually shared across an industrial facility and aligned with UTC.

When industrial 5G devices aren't actively communicating with the infrastructure, the clock domains are maintained by local clocks. These clocks gradually lose accuracy and need to be resynchronized, however. It's also possible to imagine industrial 5G devices that aren't synchronized with either a working clock or a global clock. An example is a tank sensor that sets off an alarm when the level in the tank drops too far.

In order for PTP or gPTP over 5G radio to work, 3GPP-defined device-side time-sensitive translator (DS-TT) functionality must be implemented in the industrial 5G devices. See chapter 4 for a more detailed discussion.

3.2.5 Positioning

One of the main uses of industrial 5G is for enabling the mobility of machines, materials, and people, among other things, in production and processing facilities. Mobility introduces a need for positioning.

Some HMI and xR devices involve position-dependent application behaviors. Other industrial 5G devices may also

require positioning; a gateway on an AGV, for example, can provide positioning information to help it navigate.

Another example is low-power sensors and actuators used to track materials in a factory. The device's position is reported every time it changes.

Here are some examples of positioning requirements for the analyzed use cases [6],[7]:

- When a robot is moving between two points, an accuracy of 0.3 m or better with 99.99% availability is adequate. However, it needs to reach its destination with an accuracy of ± 5 cm (this is supported by a centering station).
- When a robot is interacting with other peripherals, it may need to achieve single-millimeter precision.
- In the case of mobile tools, which have to be individually configured depending on their positions in the production line, a vertical and horizontal accuracy of better than 20 cm is required.

3.2.6 Communication Themes

The 5G industrial devices presented and described in the preceding sections require very diverse communication capabilities. Communication modules (see figure 38) and technologies linking different parts of the same device must also meet the needs of the application using it. It's therefore safe to assume that no single implementation can provide the entire range of communication parameters for all applications; aspects such as power consumption, size, and complexity can vary. On the other hand, implementing specialized modules for each profile would result in market fragmentation and make it impossible to benefit from economies of scale. Analyzing the communication requirements of various use cases and the corresponding devices, three major themes emerge.

The first is characterized by energy-efficient (battery-driven) communication and low throughput (up to a few Mbit/s; this is specified for industrial wireless sensors by 3GPP 22.104 [8]), low overall active duty with extended periods of inactivity, no essential time-sensitive data deliveries, and tolerance for temporary data loss. Devices designed for this type of situation are generally optimized for low power consumption, a small form factor, and potentially low costs. The second situation involves very high throughput (high bandwidth), low latency, and high reliability. And the third raises the bar even further with traffic-related properties that include ultra-low latency and ultra-high reliability to satisfy even the most stringent requirements of time-sensitive applications.

Communication modules and intra-device communication technologies are designed and tailored to deliver the properties that are typically associated with one of these themes. This is necessary, since the scenarios are characterized by mutually exclusive characteristics that can't all be provided by a single module. Ultimately, however, it is a product-specific decision whether or not a communication module and the corresponding building blocks for devices should be optimized to meet the needs of a particular scenario or designed to cover some of the requirements of multiple scenarios.

3.3 Examples of Industrial 5G Devices

Here we present a selection of hypothetical industrial 5G devices. They have been submitted by 5G-ACIA member companies to illustrate the possibilities going forward. None of them is available in the market at this time, and there is no guarantee that they will ever actually be developed and built.

A number of other use cases are presented in the 5G-ACIA white paper "5G for Automation in Industry" [4].

Table 2 below maps use cases and example industrial devices. It includes the use case of "portable tools", which are used throughout an assembly area to assist workers in performing specific tasks. Examples include power screwdrivers, riveting tools, and staple guns. Depending on the activity performed, they need to be configured, identified, localized, and monitored.

Table 2: Example industrial 5G devices and their potential use cases

	Motion control	Control-to-control	Mobile control panels	Mobile robots	Massive wireless sensor networks	Remote access and maintenance	Augmented reality	Closed-loop process control	Process monitoring	Plant asset management	Portable tools	Portable tools
IP67 sensor								X	X			
5G smart sensor					X			X	X			
5G IIoT level sensor									X			
5G second channel adapter									X	X		
5G remote I/O for process control								X	X	X		
Process control via mobile panel			X						X			
Mobile app for 5G industrial devices for augmented field applications							X					
5G drone operation						X						
5G Ethernet bridge	X	X		X	X			X	X	X		
5G wireless router	X	X		X				X				
5G industrial gateway				X								
5G mobile tracker					X				X	X		
5G valve terminal	X	X		X				X				
5G controller (remote I/O)		X		X				X		X		

3.3.1 5G IP67 Sensor

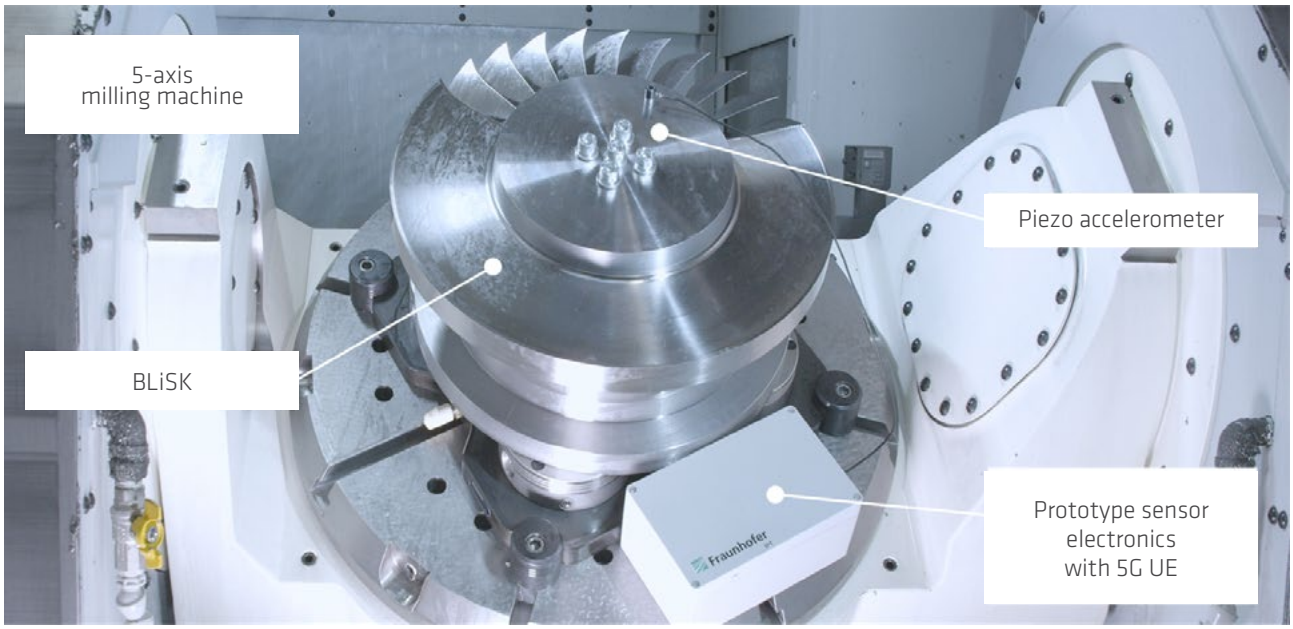
Figure 11: Example 5G sensor with integrated antennas (source: Weidmueller).

IP67 sensors are single-purpose devices surrounded by a robust enclosure for use in harsh industrial environments with varying humidity, temperature, vibrations, and other conditions. Some are also filled with epoxy resin or another insulating liquid compound to protect their internal electronics, and they have minimal external interfaces or even entirely lack them. Their power supply is often physically connected. Their main task is to reliably sense and communicate a technical process in real time, either periodically or in response to defined events. They must therefore meet exacting QoS requirements. They are purpose-optimized, cost-sensitive solutions that contain only a small number of PCB components and low-level interfaces such as SPI and UART.

3.3.2 5G Smart Sensor

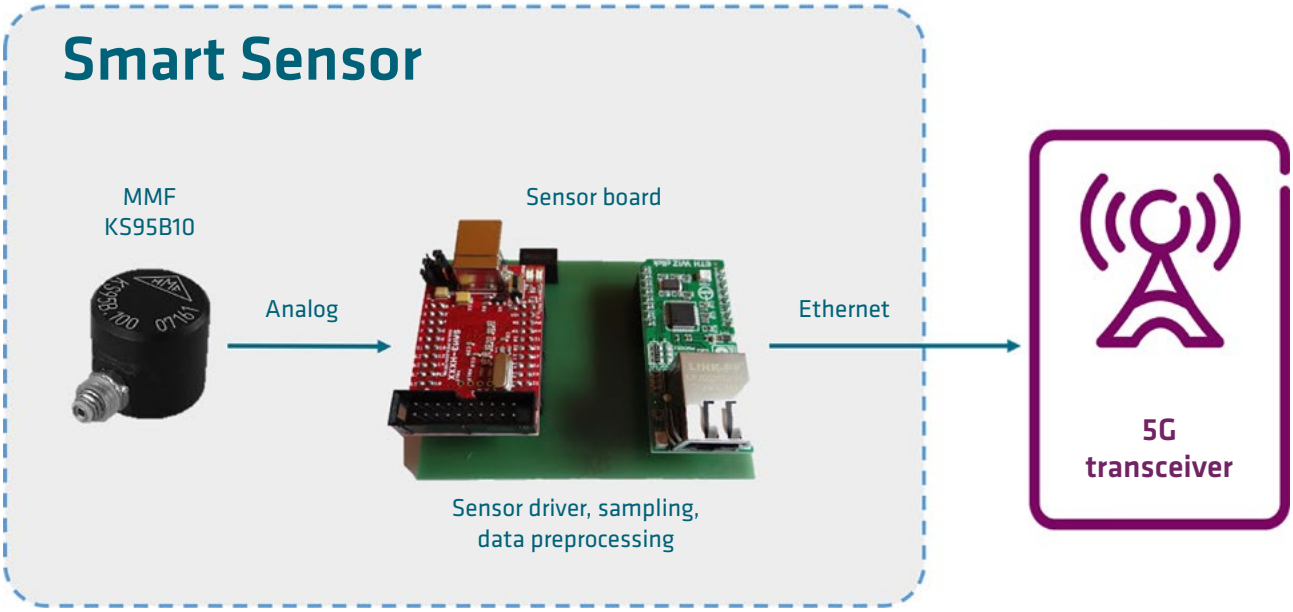
In many production applications, 5G communication lets smart sensors operate wirelessly without sacrificing reliability, availability, or low latency for short response times. Smart sensors typically have an embedded microcontroller or FPGA-based computing system for signal processing etc. While running on battery power, smart sensors can be used for machine-integrated monitoring of dynamic machining processes such as five-axis milling (see figure 12).

Figure 12: Use of a 5G smart sensor to measure acceleration in 5-axis milling (source: Fraunhofer IPT)



A smart sensor consists of a sensing probe that converts a physical quantity into an electrical signal, an A/D converter that samples the electrical signal to obtain quantified values, and a processing unit such as a microcontroller unit (MCU) or FPGA for signal processing and generation of data packets. Figure 13 shows a prototype 5G smart sensor for acceleration measurements. For 5G communication, the smart sensor can be equipped with an interface such as USB or Ethernet, linked to a 5G cellular bridge, or provided with an appropriate compact 5G communication module that is directly integrated in its PCB (once these become available).

Figure 13: Smart sensor with an accelerometer, a PCB with a sensor driver and processing unit, and an Ethernet interface (source: Fraunhofer IPT)



This smart sensor runs on battery power and can be integrated in a robust IP-grade housing (as shown in Figure 13) to allow safe operation in environments with coolants. The embedded system can be optionally used to handle different protocols such as UDP, MQTT, OPC/UA, etc. depending on the overall sensor integration concept. The sensor data can be used to trigger adjustments to the machining parameters in case any process anomalies are detected.

3.3.3 5G IIoT Level Sensor

A 5G level sensor is an example of a compact, fully integrated device for use in process industries and factories with both nonpublic standalone and public networks. Its principal task is measuring the levels of liquids or solids in mobile or fixed containers. Additional parameters, such as ambient temperature and locations, can also be detected and communicated. The device and its antenna are inside a tightly fitting enclosure (IP66/68), which restricts the possibilities for on-site commissioning and configuration. Its size is on the order of 10x10x5 cm. The device is battery-powered. The data rate is normally low (ranging from one transfer per minute down to a few per day) but may be higher (one transfer per second) when filling or emptying the container. Wireless updating of the software is also possible.

Figure 14: Mobile IIoT level sensor (source: Endress+Hauser)

3.3.4 5G Dual-Channel Adapter

This adapter connects field devices with a legacy communication protocol to a wireless 5G network. It will primarily be used for brownfield installations in the process industry to enable access to additional data for diagnosing the health of smart sensors or actuators. The adapter supports dual-channel communication (also called second communication channel in the process industry), which enables IT/OT communications independently of (wired) communication for control purposes.

The device is powered by a battery or field device and regularly transmits data at a low or moderate rate. In case there is an alarm, low-latency transmission is required. It is designed for use in harsh environments (IP66/68) including explosive atmospheres. Due to its small size of only a few centimeters across and its tight enclosure, the device doesn't include any control elements. Its antenna will preferably also be internal. It will be able to connect to both nonpublic standalone networks and public networks.

Figure 15: Field device adapter for dual-channel communication (source: Endress+Hauser)

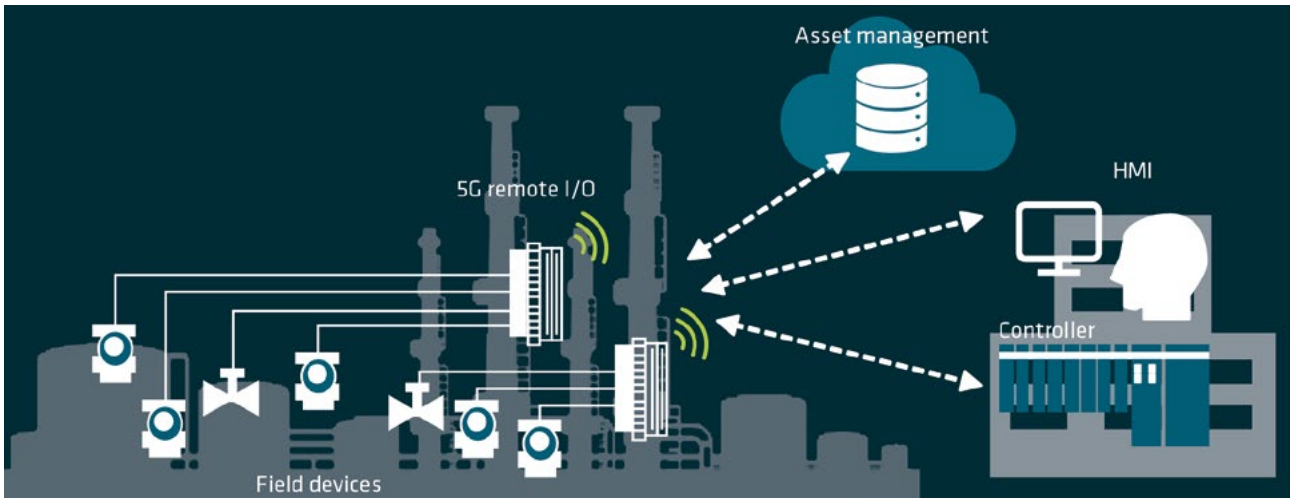
3.3.5 5G Remote I/O for Process Control

This is a modular system for linking field devices (sensors and actuators) to a plant. The devices transfer data via the I/O to and/or from upper-layer entities such as controllers, HMIs, asset management servers, etc. Reliable wireless connections may later replace the cables currently used to connect an I/O and upper-layer entities.

A 5G remote I/O is required for periodic bidirectional deterministic communication with a controller for closed-loop control, with a cycle time that is typically longer than 100 ms. The size of the messages depends on the number of devices connected to the I/O but can amount to several bytes per device. Reliable communication is critical for this use case; if it is lost, the entire plant can stop functioning. The requirements in terms of the spacing and reliability of messaging can be relaxed for process monitoring purposes. The 5G remote I/O can also carry noncritical data for device management operations such as diagnostics and software updates as required by the operator.

A 5G remote I/O is required for operating reliably and safely in harsh environments (e.g. across a temperature range from -40° to 70°C and relative humidity between 5% and 95%), including zone 2 hazardous areas. It is stationary and receives its power supply from an external source via a cable.

Figure 16: 5G remote I/O installed in plant field (source: Yokogawa)

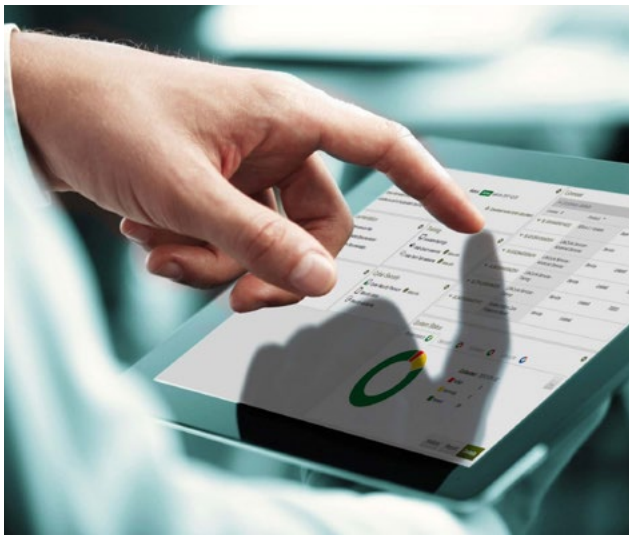


3.3.6 5G Process Control via Mobile Panel

A battery-operated mobile panel gives a plant's operators and workers instant access to the production environment, letting them monitor and control the status and setpoints of processes from any location within the plant. Operator mobility within a facility can be provided by using 5G for connectivity between the mobile panel device and distributed control system.

The device displays information from the distributed control system and lets users take action while on the shop floor to concurrently supervise multiple automated processes. It allows them to "see what it sees", thus reducing the time needed to optimize a process or correct a problem.

Figure 17: Process view using a mobile panel (source: ABB)



3.3.7 Mobile App for 5G Industrial Devices for Augmented Field Applications

Figure 18: Smart glasses for process monitoring (source: Endress+Hauser)

This mobile 5G device with an app supports augmented field applications to improve how work is done to a greater extent than what is possible with conventional paper-based approaches. The operator gets an up-to-date view of scheduled tasks and step-by-step support for executing procedures. The solution eliminates confusion about which is the latest version and facilitates updating, copying, and distribution of it to relevant personnel. It also provides the operator with knowledge management tools, including easy access to additional information (pictures and manuals). Operators can use a built-in camera to take pictures of the steps involved in procedures or read QR codes to ensure that work is executed using the correct equipment. This enables operators to acquire greater competency while performing tasks.

Industrial 5G devices such as tablets, mobile phones, edge gateways, and smart glasses can significantly improve the end user experience with augmented reality (AR) features.

5G-enabled mobile field workers using an augmented field procedure need the mobile app to integrate control system data and context- and situation-awareness functions. This way they can receive field information in real time, automatically capture values, and directly interact with any control system to execute procedures in a synchronized manner. This improves the efficiency of work and reduces the need for control room and field operators to constantly communicate with one another by radio. The solution provides relevant instructions while helping to ensure that work is done correctly. It also includes other features to support field workers, such as voice synthesis, remote assistance, and an industrial chatbot.

This opens up the possibility of remotely executing process control actions over the 5G network (for example, opening or closing valves). It therefore requires a time-synchronized network in which packets are received on time and in the right sequence. Data transmitted over the 5G network will need to be timestamped on the device and network levels.

3.3.8 5G Drone Operation

Many of the unmanned aerial vehicles (commonly known as drones) in use today are controlled by a human operator via a point-to-point link over a private wireless network or ISM band.

5G-enabled drones can significantly improve the user experience by using a public or private (nonpublic) 5G network for monitoring large and distant areas with high-performance communication. Such a drone is equipped with sensors (e.g. an IR sensor) for fast, efficient monitoring, surveillance, and inspection of areas such as industrial sites. The captured sensor data is continuously relayed to the user for further analysis.

For such a 5G-enabled drone to operate reliably, the following would be required:

- A control system characterized by high availability and security and low latency. Real-time positioning and time synchronization capabilities are also a must.
- The data captured by sensors installed on the drone has to be sent to the user over the network, which requires a high uplink throughput.

5G-enabled drones will be battery-powered and have an appropriate IP rating for outdoor operation. Figure 19 shows an example.

Figure 19: Drone operations via 5G (source: ABB)

3.3.9 5G Ethernet Bridge

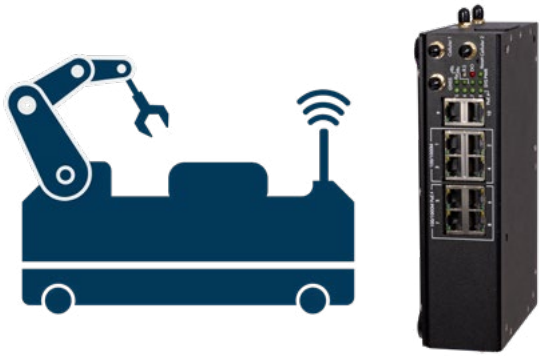
Figure 20: 5G Ethernet bridge in a typical cable replacement use case (source: HMS Networks)

A 5G Ethernet bridge can be used to link Ethernet devices to a 5G network. This is typically done to replace cables with a wireless solution as illustrated in figure 20. In this use case, an industrial Ethernet protocol is bridged via the 5G network. It is also possible to use this kind of device for normal IP traffic.

The prerequisites for this to work are high reliability, low latency, and accurate time synchronization. The device has IP65 ingress protection and internal antennas. The required data throughput is typically less than one Mbit/s. One exception is when the device is used for 2D or 3D sensors; in this case, an uplink speed of several hundred Mbit/s is required. The device is often powered by a battery in a mobile machine, in which case its mobility is limited to that of the machine.

3.3.10 5G Wireless Router

Figure 21: A mobile machine and a 5G wireless router (source: HMS Networks)



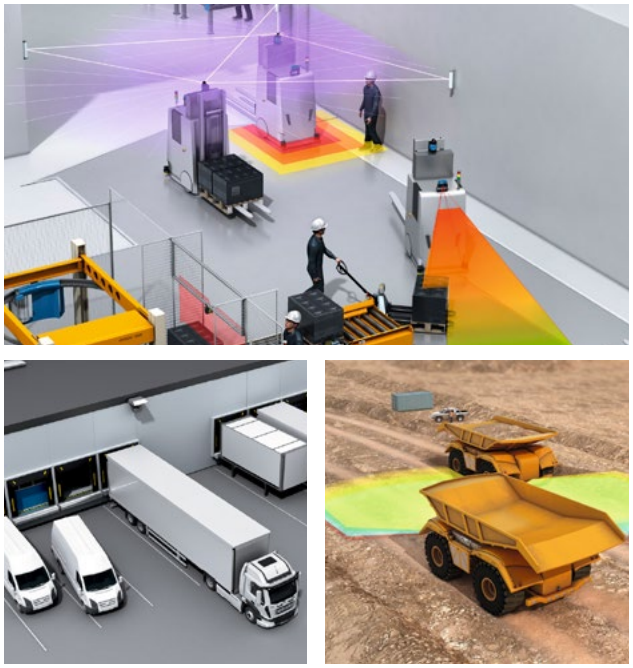
Here a 5G wireless router doubles as a LAN switch for connected devices. A typical use case is mobile machine connectivity as illustrated in figure 21. The mobile machine is used in conjunction with a traffic management system and other IT functions. The mobile machine can optionally also use a safety protocol.

This requirements for the device to communicate with the router are high reliability, low latency, and accurate time synchronization. It is IP30-rated and equipped with external antennas. The required throughput is typically less than one Mbit/s except when used for 2D or 3D sensors, in which case an uplink speed of several hundred Mbit/s is needed.

The device is often powered by a battery on the mobile machine and the mobility of the device will be determined by the mobility of the mobile machine.

3.3.11 5G Industrial Gateway

Figure 22: Indoor industrial vehicles and outdoor automation (source: SICK AG)



A modular 5G industrial gateway can be used for indoor industrial vehicles and mobile outdoor automation for machine to machine, machine to infrastructure, and machine to fleet manager communication.

Currently, tasks such as localization, personal safety, collision protection, and load handling are mainly solved locally on each vehicle with only minimal communication with its environment. This limits the efficiency of indoor industrial vehicles and rules out the possibility of automating machines that are used outdoors. However, a new solution integrates detection and identification systems in the active vehicle, reliable wireless communication with other machines, infrastructure-based environmental monitoring with various sensor technologies, and continuous reporting of environmental data that can also be used to update maps and optimize routes.

A 5G industrial gateway can carry both **cyclical data** (for safety-related applications, at approx. 100 kbit/s with a cycle time of less than 100 ms) and **noncyclical data** (for example, transferring data for map updates in bursts at a speed greater than five Mbit/s). In special cases, sensors or cameras may send raw data from machine to machine or to an edge computer. Time synchronization is needed for these scenarios.

Especially outdoors, sidelink communication between devices can be crucial for compensating for coverage gaps in 5G system antennas.

Ubiquitous positioning with roughly 0.5-meter accuracy (using GNSS or 5GS) that could be refined further using other positioning techniques at loading/unloading.

Powered by the vehicle (if its engine is running and/or it has a large battery), so energy consumption isn't a critical factor. When the machine isn't operating, it can go into a low-power mode for tracking purposes.

3.3.12 5G Mobile Tracker

Figure 23: Uses for mobile indoor and outdoor trackers (source: SICK AG)

This application involves a battery-powered tracker containing a 5G communication module, along with integrated sensors for condition monitoring and tracking the locations of transported goods, objects in ports or airports, non-power tools, and waste/fill level management. The use case scenarios for devices of this kind pose different requirements with regard to the form factor and IP class. The security and authentication methods used should be suitable for low-complexity IoT devices. The device will typically send data to an (edge) cloud.

Use cases involving location and mobile tracking must consider variables such as international reach, regulations, density requirements, and consistency across indoor and outdoor environments.

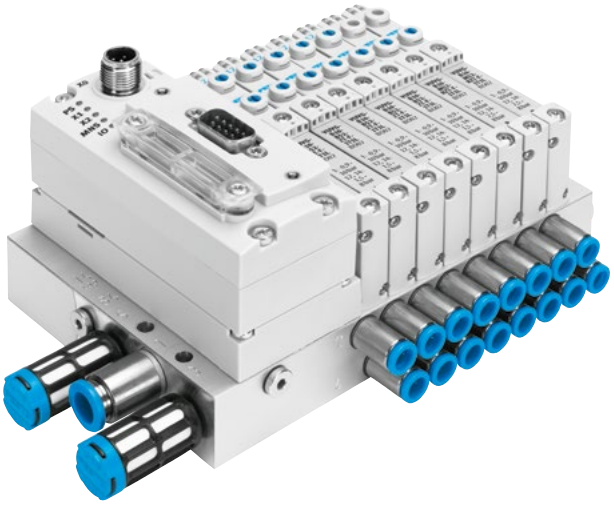
The device sends a small volume of data (at a rate of around 100 kbit/s) in a burst lasting several seconds. This can be triggered by an event or the elapse of a defined time interval. It is not intended for continuous monitoring, for instance of an engine's vibrations. A deep sleep mode can be used to save energy, but its range of possible uses is limited by the lack of a way to wake it up again remotely.

There is a need for an indoor and outdoor (low-energy) positioning capability with an accuracy of between five and 100 meters using 5GS, GNSS, or another wireless technology such as Wi-Fi or BLE. The actually required accuracy will depend on the use case and situation.

Low power is critical for enabling long recharging and/or battery replacement cycles, for example 13 months apart with batteries being replaced within the scope of yearly maintenance.

3.3.13 5G Valve Terminal

Figure 24: 5G valve terminal (source: Festo SE & Co KG)



A valve terminal is mainly used to operate multiple channels in pneumatically controlled systems without the need for a switch cabinet. Its modular mechanical design integrates multiple pneumatic valves and a controller for decentralized control tasks.

An integrated microcontroller provides processing capabilities as part of the integrated control unit. Interfaces for sensors and diagnostic data enhance the terminal's functionality.

Legacy fieldbuses and industrial Ethernet are established technologies for communicating with higher-order PLCs. 5G URLLC will replace these wired connections and deliver additional benefits for flexible production plants.

Typical use cases with challenging timing requirements include robotic front ends, potentially also in moving applications.

3.3.14 5G Controller (Remote I/O)

Figure 25: Shown here is the WAGO PFC200 4G controller; a 5G device could be similar to it (source: WAGO GmbH & Co. KG)

A 5G controller resembling this one could be used to pre-process data from sensor and actuators or provide access to these peripherals as a remote I/O in a wireless network. Common digital or analog sensor and actuators, which don't need to have IP-based communication, could be directly wired to I/O modules for flexible connection to the controller. The typical use cases include controlling flexible machine parts in a control-to-control loop, collecting data for energy data management at large production sites, and controlling applications installed on an AGV.

Depending on the use case, this device requires low latency and high reliability. High data rates aren't necessary; one Mbit/s is normally sufficient. Faster data rates are useful for software updates but don't need to exceed 10 Mbit/s. This device is intended for installation and use inside a box and therefore doesn't need to be designed to withstand harsh environments on its own. It doesn't include a battery but can be used on a mobile machine because of its low power consumption. The 5G controller has no internal antennas and can be used with both public and nonpublic standalone networks.

4. Logical Reference Architecture for Industrial 5G Devices

To shed light on how the various components of a device implementation interface with one another, this chapter presents reference architectures in the form of generic block diagrams for the most common types of industrial 5G devices. There are many ways to do this, depending on which of a device's physical resources meet the requirements of which logical functions. We start with a generalized, undifferentiated logical architecture.

An industrial 5G device's logical architecture depicts what it does without considering the actual hardware components used to implement it. It shows the device's main functions from both the ICT and the OT perspectives and how they are supposed to interact with one another. Once this has been done, it is easier to progress to a block diagram for implementing the detailed architecture.

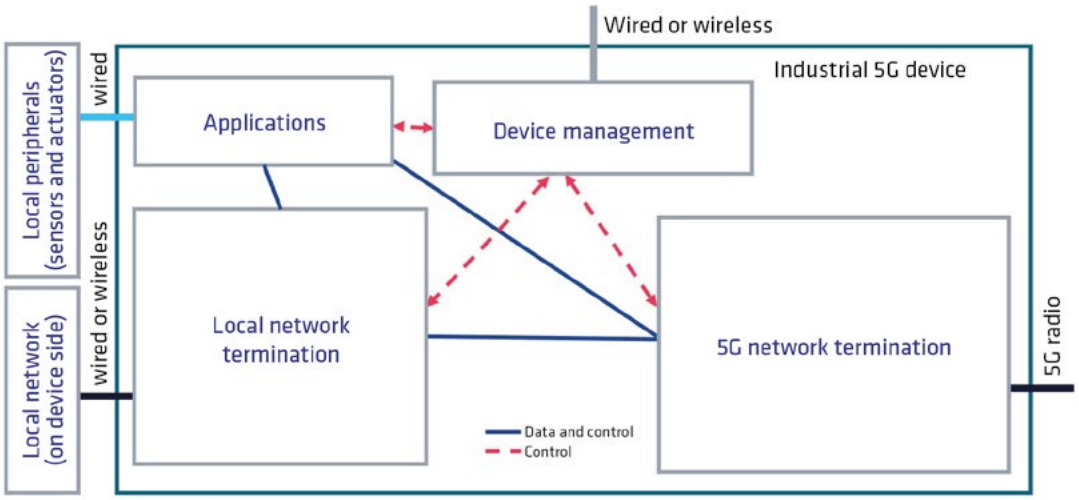
The architecture integrates functions that aren't always present in mainstream 5G devices but are important for industrial devices. They include Ethernet bridging, IEEE Time Sensitive Networking (TSN), and capabilities related to Precision Time Protocol (PTP), all of which are important for devices operating in industrial Ethernet- or IP-based networks. They also include EAP-based authentication, which is relevant for devices operating in nonpublic 5G networks.

This section starts by looking at the top-level functional architecture and then goes on to describe each top-level function in enough detail to ascertain the interfacing requirements for the implementation-level block diagram architecture.

4.1 Top-Level Logical Architecture

An industrial 5G device is a managed connectivity device whose main purpose is to provide 5G connectivity for one or more applications or other devices serving an OT operation. The applications can be integrated in the device itself or connected to it via a local network. Figure 26 shows a top-level logical architecture derived from this functional basis. This architecture includes all top-level functions that would be needed in at least one type of industrial 5G device (not all of them would be needed in every type).

Figure 26: Top-level logical architecture



The top-level architecture focuses on communication capabilities and contains the following elements:

5G network termination
This comprises all 3GPP-defined device-side functions for connecting to a 5G network and operating as part of it. This function makes an industrial device 5G-enabled and is a requirement for all types of devices.

Local network termination
This is required in order for an industrial 5G device to connect to a local network (provided that it also has a local network interface).

Device management
This function is included in the top-level functional architecture, based on the assumption that there will be a need to manage the industrial 5G device's 5G- and OT-related functions. For the sake of simplicity, a set of management functions is depicted in the top-level architecture as a single generalized function.

Applications
These comprise all higher-layer functions residing inside the industrial 5G device that aren't covered by any of the other top-level functions. Besides measurement and automation functions, these applications also include functions for the sensors and actuators that are integrated in the device and/or interface functions for peripheral sensors and actuators. HMI with xR devices may also include communication media for interacting with humans (such as video screens, cameras, loudspeakers, and microphones).

The blue lines in the middle of the diagram connect the 5G termination with either or both of the OT functions supporting the application within the device or local network termination, and convey both payload data and associated control signals. The dashed red arrows leading from the common and shared functions to all other functions represent control signaling paths.

While defining the device's logical architecture, the main focus is on understanding its internal composition and interconnections. However, external interfaces can also be important for an industrial 5G device's overall functionality. As a minimum, every type of industrial 5G device must have a 5G radio interface, and may also optionally have a local configuration interface as shown at the top of the figure. If an OT application residing inside the device relies on external peripherals such as sensors and actuators, one or more external point-to-point interfaces are needed to support this, as shown at the top left. If some of the applications need to be reached via a local OT network, a networking interface (shown on the bottom left) is also required.

4.2 Practical Logical Architecture

It's important to present the architectural details down to the level at which the blocks and interfaces of the implementation architecture come into view. It makes little sense to break them down any further than this, since they are either likely to be implemented inside a single component or it is clear that the functions concerned will only be implemented on the device's OT or 5G side without any other interfaces. There are many types of OT functions and applications, for example, but it would exceed the scope of this white paper to cover all of them. Here it's enough to highlight the different kinds of communication requirements that can apply in a 5G context and provide a few examples of different types of industrial 5G devices.

As already discussed, one important architectural aspect is whether a device integrates the application that serves its OT functions and whether or not it is able to connect to a local network. Also important is the extent to which applications require support for QoS and time synchronization.

Considering these aspects and the industrial devices introduced in chapter 3, four different logical architectures enter into consideration. These are introduced here and described in greater detail in the following sections:

- The first kind of logical architecture (section 4.2.1) involves a type of device that directly hosts all required OT applications. It isn't connected to any local networks on the device side and therefore doesn't need to include a local network termination function.
- The second kind (section 4.2.2) is enhanced by local network termination capabilities. It involves devices that can serve as either an IP host or router or an Ethernet end station, bridge, or application-layer gateway. They are appropriate for applications or networking scenarios that only require conventional IP and Ethernet quality of service (DiffServ, Ethernet

traffic classes) and don't rely on support from either IEEE TSN traffic scheduling or shaping functions or accurate PTP time synchronization over a 5G radio link. In practice, this means that the device doesn't need to include any device-side time-sensitive networking translator (DS-TT) functionality as defined in 3GPP releases 16 and 17.

- The third kind (section 4.2.3) refers to a device that additionally supports accurate (g)PTP-based time synchronization (according to IEEE 1588 and/or IEEE 802.1AS) over 5G radio. For this purpose, the device needs to implement a subset of DS-TT functionality that is relevant to (g)PTP as defined in 3GPP release 17. It doesn't necessarily need to include full IEEE TSN-capable DS-TT as defined in 3GPP release 16. If the device has (g)PTP specific DS-TT capabilities, it may operate either as part of the 5GS bridge or as a standalone Ethernet bridge or IP router, depending on the 5G network capabilities and overall network setup. This device architecture and these capabilities are suitable for deployment scenarios in which conventional IP or Ethernet QoS with accurate PTP time synchronization is adequate and neither IEEE TSN traffic shaping nor scheduling is used.
- The fourth (section 4.2.4) occurs in devices that also need to be able to operate as part of a 5GS bridge that supports the IEEE TSN-compliant centralized configuration model with IEEE TSN functionality that was introduced in 3GPP Release 16 and augmented in Release 17. This requires the device to include DS-TT that specifically supports the IEEE 802.1AS PTP profile used in IEEE TSN and the LLDP protocol used for Ethernet topology discovery.

Including a detailed list of PTP- or TSN-related features and profiles would exceed the scope of this white paper, which takes an architectural perspective.

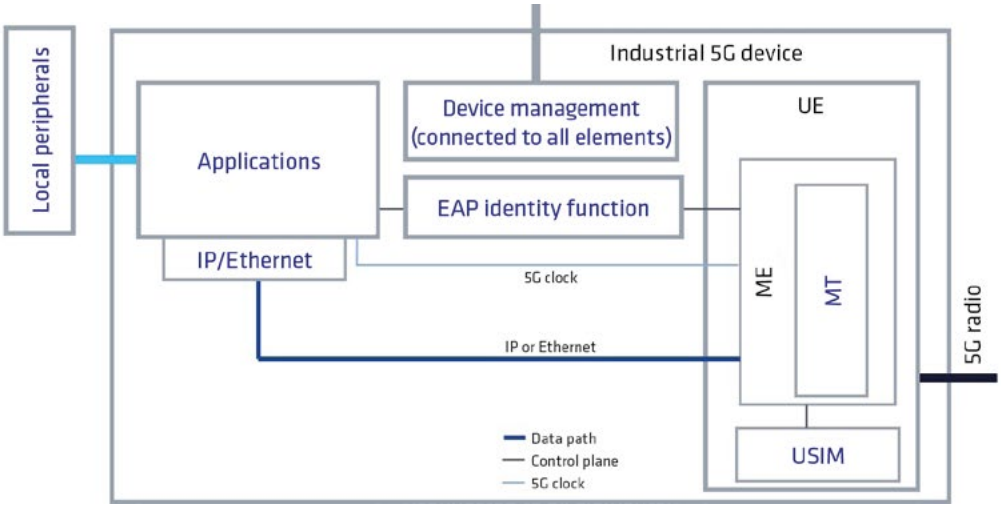
Another consideration for detailed work is that, while the logical functions are independent of the actual implementation, it is helpful to acknowledge the implementation technologies that are clearly going to be used in any case, like Ethernet-based technologies on the device-side local network. Ethernet PHY is therefore included in the logical architecture schemes shown below for that interface. Please note that while some local peripheral interfaces will also use Ethernet-based technologies, it can't be assumed that this will generally be the case, so no technology label is applied to that interface.

4.2.1 Logical Architecture for Supporting Applications Inside a 5G Industrial Device

In the logical architecture shown in figure 27, OT applications – like those serving sensors or actuators – are either embedded in the device itself or, as shown in the figure, connected as local peripherals. This is an architecture that doesn't connect to a local network on the device side and therefore doesn't need a local network termination function.

The device can have an application layer gateway function between its Ethernet or IP connectivity on the 5G network side and use any protocol or technology to link to local peripherals. See figure 8 c) and d) for examples.

Figure 27: Logical architecture for supporting applications inside a 5G industrial device



The top-level "5G network termination" function discussed above mainly maps to the UE shown in figure 27, which consists of a mobile equipment (ME), a mobile termination (MT), and a universal subscriber identity module (USIM). The last of these also hosts 3GPP-based AKA authentication functions. The USIM is a logical function hosted on the universal integrated circuit card (UICC). For visualizing the internal interfaces of an industrial 5G device, it's enough to keep in mind that all of the connections to the UE terminate at the ME.

Extensible authentication protocol (EAP) authentication, which is hosted by the extensible authentication protocol identity function (EIF), can be used for connecting to nonpublic networks. It's connected to the ME for control signaling (indicated by a gray line). EIF is needed to apply authentication methods other than EAP authentication and key agreement (EAP-AKA) when no suitable USIM is available. This function isn't covered by the 3GPP standards. EIF is functionally similar to USIM in the sense that it can also be used to store subscription information and security credentials and also terminates the EAP protocol. The possibility has already been discussed that it can also be necessary to connect the EIF to the application hosting the OT-related functions in the device. This control signaling functionality is therefore included in the figure as an option. Since EIF is used with OT networks, it's important to enable flexible OT-defined deployment and provisioning options for it. Depending on the network's requirements, either USIM or EIF can be used as the primary authentication instance. The industrial 5G device can support both.

This logical architecture has a "5G clock" interface between the 5G communication module (ME) and the applications. The ME can be synchronized with 5G time (typically traceable to UTC), and the interface lets 5G time be distributed to local applications as well. This is a basic time management capability that lets applications use 5G time for timestamping events (such as a measurement made by a peripheral) and also comes into play when a global time domain is needed for subsequent processing of measurements (for example, to determine the order in which events have taken place).

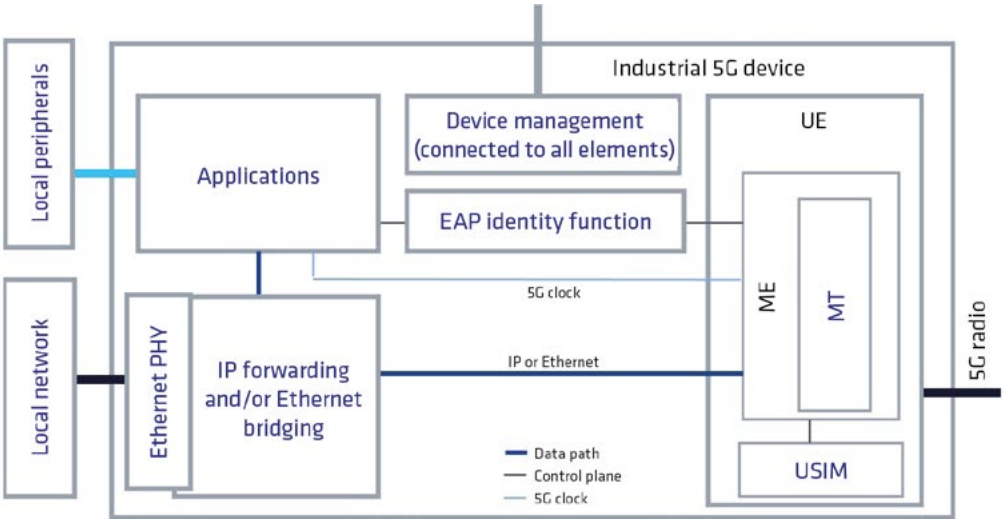
Depending on the needs of applications and the network or other end stations to which a device is connected, it may be necessary to provide IP and Ethernet protocol functionality with more advanced time management capabilities and support for QoS. The following sections (4.2.2, 4.2.3, and 4.2.4) describe different versions of these capabilities for the local network termination function used to connect a device to a local network. They can also be used for applications inside a device.

It is assumed here that all of the logical elements shown for the logical architectures described in section 4.2 need to be managed, either separately or together with other functions. The logical architecture includes a device management function that can be contacted via a local or remote management interface and is able to execute management actions for all of the device's functions. This calls for logical links to all of the depicted functions. For remote management, the device management function can present itself as an Ethernet or IP-based application inside the device for sending and receiving remote management commands, which are treated as payload traffic in the local or 5G network. For simplicity's sake, these connections have been omitted from the figure.

4.2.2 Logical Architecture for Supporting Applications or Networking Using IP or Ethernet with Traditional Non-Time-Aware QoS

Figure 28 shows the logical architecture for cases that require support for QoS but not for TSC/TSN. 5G connectivity is used to support Ethernet or IP traffic. Both are shown here: traffic to and from the application integrated in the device, and traffic routed to the device's Ethernet port, shown by the blue line in the middle. This line only represents traffic that the 5G system recognizes as user payload traffic; there is no direct interface for control signaling between these elements.

Figure 28: Logical architecture for applications using IP or Ethernet with QoS support



This logical architecture contains the same functions as the logical architecture that supports applications inside the 5G industrial device (see section 4.2.1) while adding functions related to local network termination. This enables the device to execute an Ethernet bridge function as shown in figure 8 a), an IP router as shown in figure 8 b), or an application-layer gateway (implemented as an application within the device) between IP- and/or Ethernet-based application protocols. The Ethernet network bridge and IP router can support QoS via mechanisms such as DiffServ or Ethernet priority code points (PCP) while mapping them to 5G QoS on the 5G network side.

The 5G network termination (UE), application, EAP identity function (EIF), 5G clock interface, and device management are identical to those already described in section 4.2.1.

In the context of this logical architecture, 5G time could also be distributed to the local network connected to the 5G industrial device by NTP, PTP, or some other method. However, this kind of logical architecture isn't suited for accurately distributing external time domains via PTP over 5G radio, which introduces jitter. Synchronization to and distribution of an external working clock signal via PTP requires the capabilities provided by the 3GPP device-side time-sensitive networking translator (DS-TT) function, which is the central element of the corresponding logical architecture described in section 4.2.3.

4.2.3 Logical Architecture for Supporting Applications Using IP and Ethernet with QoS and Precision Time Protocol over a 5G Radio Link

Time synchronization is important for many industrial applications. 3GPP has defined a set of functions for supporting IEEE TSN; it is applicable to applications that have been specifically designed for TSN. 3GPP has specified that these functions must reside in a device-side time-sensitive networking translator (DS-TT). However, many network deployments and use cases don't require the full set of TSN traffic scheduling or shaping-related features; support for accurate PTP time synchronization is sufficient in conjunction with conventional IP and Ethernet QoS mechanisms. For these purposes, 3GPP Release 17 will include the possibility of having a DS-TT with only PTP-specific capabilities.

To sum up, the DS-TT is needed to deduce exactly how much time a PTP (sync) message has spent inside the 5G system (called the residence time), in other words between the DS-TT and the network-side TSN translator (NW-TT), which acts similarly to the DS-TT in the 5G core network user plane function (UPF). 3GPP Release 16 requires the PTP grand master clock to be on the UPF/NW-TT side of the 5GS, with PTP sync messages only being delivered to devices/DS-TTs in the downlink direction. Release 17 also allows PTP GM on the device/DS-TT side with delivery of sync messages in the downlink direction and, via the UPF, to other devices/DS-TTs.

It is possible to determine the residence time between NW-TT and any DS-TT or between two DS-TTs because NW-TT and all DS-TTs are synchronized with 5G time. The time-sensitive networking translator on the egress side, which is either a DS-TT or a NW-TT depending on the direction, inserts the residence time value into the PTP packet headers as a correction term. The NW-TT and DS-TT operations for PTP are necessary when time synchronization accuracy on the order of microseconds is required, owing to the variable delay introduced by 5G radio.

Generally speaking, a device with DS-TT that supports (g)PTP but not TSN traffic scheduling capabilities can be used in two types of network deployment scenarios:

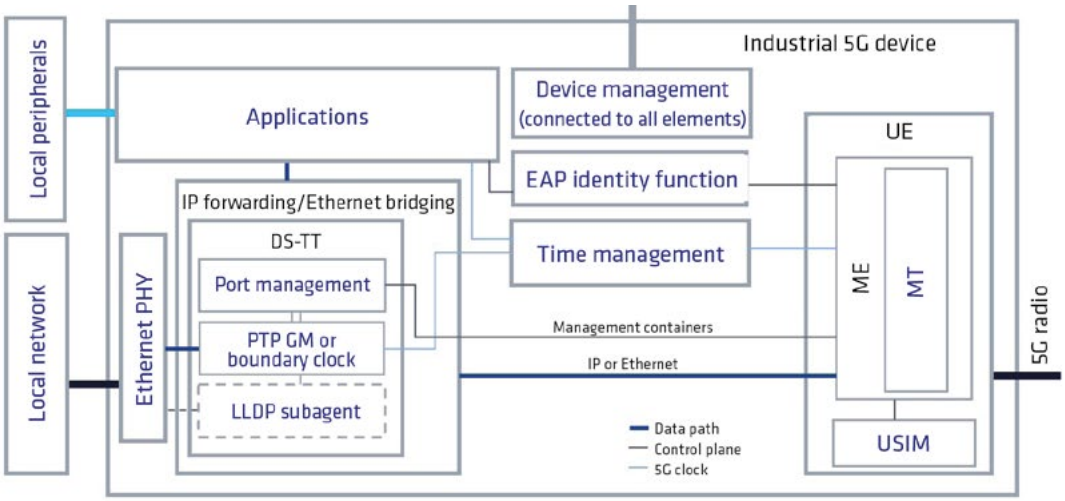
1) The device is connected to the 5G network using an Ethernet PDU session and acts as a port in a 5GS bridge formed by 5G UPF/NW-TT and other devices. The bridge can operate as an IEEE 802.1AS (gPTP profile) time-aware system or as an IEEE 1588 (PTP) boundary clock or transparent clock. The port, including its (g)PTP operation, is managed by a special 3GPP-specified TSN application function. The 5GS bridge as a whole, modeled as a PTP instance, may operate and be managed as part of an (industrial) Ethernet network when (g)PTP support is required. Support for gPTP profiles is specified in 3GPP Release 16, while the other type of PTP support is specified in Release 17.

2) The device is connected to the 5G network using an Ethernet or IP PDU session and, along with UPF/NW-TT and possibly other devices, modeled as a PTP instance that can work as a IEEE Std 802.1AS time-aware system (for Ethernet only) or as an IEEE Std 1588 boundary clock or transparent clock. Operation of PTP instances can be managed by any application function using the 3GPP NEF time synchronization API. This deployment scenario with 5G-managed PTP operation is specified in 3GPP Release 17.

The DS-TT is managed via management containers carried in the 5G control plane, so the DS-TT also needs the ability to send and receive them. This is shown in the figure as a special "management containers" interface.

This logical architecture version is shown in figure 29 below.

Figure 29: Logical architecture for supporting applications via IP and Ethernet with QoS and PTP time synchronization



The 5G network termination (UE), the application, the EAP identity function (EIF), and device management are the same as those described for the first logical architecture version in section 4.2.1. This version also has other capabilities, since it incorporates functions related to time-sensitive networking.

A time management function has been introduced for interlinking multiple clocks running in the device, in case a simple point-to-point interface isn't sufficient. The time management function is also a basic feature of functions that are needed to manage different time domains within a device, and a connection to the application is therefore also included here as an option.

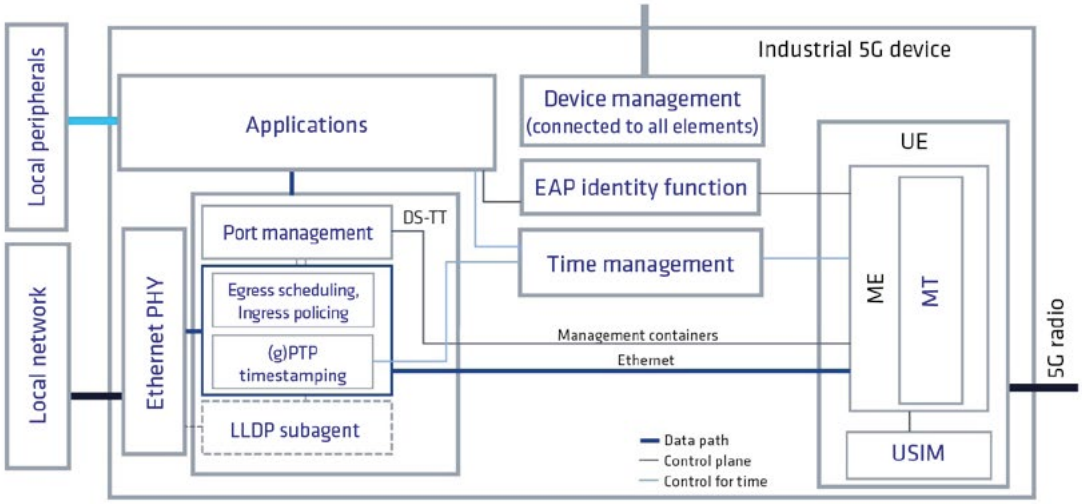
The DS-TT can also optionally include support for the IEEE link layer discovery protocol (LLDP). LLDP (IEEE 802.1AB) is used for topology discovery for Ethernet. It is mandatory for TSN-capable bridges, but can also be used in non-TSN-specific Ethernet deployments.

4.2.4 Logical Architecture for Supporting Applications Using Ethernet with IEEE TSN

Figure 30 shows a logical architecture version containing the functions described above as well as all of the other functions that support TSN, in particular the time-aware shaper (IEEE 802.1Qbv). This lets an industrial 5G device act as a port for TSN-capable 5GS bridges as defined in 3GPP Releases 16 and 17. A 5G industrial device acting as a standalone Ethernet bridge is unable to support TSN traffic shaping or scheduling features over 5G radio on its own due to the delay variability of the radio; it has to join the distributed 5GS bridge for this purpose.

Where PTP is concerned, TSN requires support for the IEEE 802.1AS profile of PTP, for which the DS-TT functionality is described in detail in 3GPP Release 16. This logical architecture may be considered to be the most advanced version.

Figure 30: Logical architecture for applications using Ethernet with IEEE TSN



From the 5G network termination perspective, this version differs mainly in its ability to control user plane payload traffic while interfacing the local network with time-sensitive network (TSN) scheduling. For this purpose, it integrates a DS-TT function defined by 3GPP that includes egress scheduling and ingress policing. The DS-TT scheduling and policing parameters are configured by the time-sensitive networking function via the same management containers that are used for PTP management. This function exposes management of the entire 5GS bridge via standardized IEEE interfaces to the centralized network controller (CNC) for time-sensitive networking, which is ultimately what provides the DS-TT scheduling and shaping configurations.

The 5GS bridge and TSN both require the 5G core network to support TSN/TSC capabilities as defined in 3GPP Release 16 and enhanced in Release 17. The network also needs to deploy the time-sensitive networking function for bridge management.

4.3 Device Authentication

4.3.1 Introduction

Mutual authentication between a 5G network and 5G device is based on the conventional USIM model known from previous cellular generations. Summing up, the USIM holds the device's permanent 5G-specific identity (referred to as the SUPI in the context of 5G), a long-term secret key shared by the USIM and network, and a cryptographic algorithm that permits mutual proof of possession of this long-term key. The USIM also stores the subscriber profile, which includes but is not limited to network-specific cellular parameters that define how the device behaves toward a given network (a list of preferred networks is one example).

Authentication is executed between USIM and the network during initial network registration (in other words, when a cellular device attempts to connect to the network). This is referred to as primary authentication.

3GPP defines two types of industrial networks that take different approaches to device authentication and the Universal Subscriber Identity Model (USIM); they are described in the following two sections. In the first approach, private networks piggyback onto the infrastructure of a public mobile network (the PNI-NPN scenario), while in the second private networks don't rely on the functions of a public land mobile network or PLMN (the SNPN scenario).

Besides primary authentication, 5G includes the concepts of slice-specific authentication and authorization, which take place completely independently of a USIM. They are covered in section 4.3.4 below.

4.3.2 Primary Authentication for PNI-NPNs

In network deployments that take the public network interface nonpublic network (PNI-NPN) approach, one or more slices or cells (constituting "closed access groups") of the public network are dedicated to a specific OT network. The device must use a USIM issued by the public mobile operator in order to attach to the PLMN network that is providing the resources for the PNI-NPN.

Since an ordinary (IMSI-based) operator USIM is involved, its deployment is bound to UICCs (or eUICCs), and mobile operator procedures are applied to distribute and manage it.

In the case of PLMNs (including PNI-NPNs), it's mandatory to deploy a USIM (universal subscriber identity module) on a dedicated secure element called a UICC (universal integrated circuit card). In the context of remote provisioning, there is no such thing as a USIM permanently coupled with a UICC. For previous cellular generations, GSMA had already introduced the possibility of dynamically deploying USIM profiles (a text description of the entire content of a USIM) as embedded universal integrated circuit cards or eUICCs. The main difference between an eUICC and a UICC (which also exists in soldered form) is the possibility of storing USIM profiles in the eUICC.

The geometry of a USIM deployment is clearly relevant to the physical layout of an industrial device, and it has an even greater impact on how industry verticals use key management and distribution procedures.

Physically distributing and inserting removable cards could work well in a limited number of entry scenarios, but isn't an economically viable option for complex, large-scale deployments. Capabilities for electronically deploying ("provisioning") USIM profiles in a device are essential, however.

One option for SNPNs is to adopt eUICCs and the GSMA's remote SIM provisioning framework. However, although administrative issues related to certification requirements could definitely be resolved, it is unclear whether this approach could provide optimal synergies between existing key and identity management approaches at OT companies and management of 5G-specific identities and credentials.

The next section therefore goes into detail on how primary authentication can be executed with SNPNs without having to rely on USIMs and UICCs.

4.3.3 Primary Authentication of SNPNs

Since a SNPN doesn't rely on network functions provided by a PLMN, the corresponding 3GPP specification [TS 33.501] allows for the use of new primary authentication methods (apart from USIM-based ones). The choice of (EAP) authentication methods is left to the private network owner, for example the OT operator. How identities and credentials for these new methods are stored and processed in a device is beyond the scope of the 3GPP specifications. This paves the way for industrial devices without UICCs or eUICCs.

Both aspects are discussed in greater detail below.

New EAP Methods

Up to Release 15, the only available authentication method was the AKA (authentication and key agreement) protocol, which uses symmetric keys shared by the USIM and network. Two variants of AKA exist within 5G: 5G AKA and EAP-AKA'. 5G AKA has evolved from the EPS-AKA protocol used for previous cellular generations, while EAP-AKA' is an adaptation of the AKA protocol used with the extensible authentication protocol (EAP).

3GPP has added the EAP framework to enable new authentication methods that could be especially helpful for industry verticals in private networking scenarios. One example, which is expected to be relevant to industrial deployments, is the EAP-TLS protocol; it uses private public key cryptography instead of shared symmetric keys. In addition to specifying in detail the implementation of EAP-TLS for 5G authentication, Release 16 has introduced a new type of permanent identifier as an alternative to the conventional IMSI-based SUPI consisting only of decimal digits. A SUPI of this new network-specific identifier (NSI) type has the form <username>@<realm>.

While the EAP framework and new SUPI type were important steps toward authentication schemes optimized for verticals, 3GPP has left open how EAP should be implemented in devices.

In the current USIM architecture, the EAP protocol is terminated by the mobile equipment, in other words outside the USIM. During authentication, the USIM is invoked for a cryptographic operation that uses a single command and is the same for both EAP-AKA' and 5G AKA.

This command can't currently be used for EAP-TLS. 3GPP would have to define new commands and storage capabilities for a private key and certify the USIM in order for it to be used in combination with an USIM. It should also be noted that SUPIs of the new NSI type can so far only be used in combination with AKA protocols, although 3GPP has specified a dedicated non-IMSI variant of the USIM in Release 16.

Migrating the EAP client of mobile equipment to a new authentication client while adding generic EAP support may provide benefits by making it possible to introduce new EAP variants without modifying the equipment. In the context of this white paper, the term EAP identity function (EIF) is proposed for designating such an EAP-enabled authentication client (replacing the USIM), despite the fact that no formal specifications exist for it.

Primary Authentication Without UICC

One property of the EIF is that (as opposed to USIMs), 3GPP doesn't define any requirements (such as use of an UICC) related to its deployment. The industrial 5G device's manufacturer may therefore implement the EIF in accordance with the requirements of a particular industrial use case, for instance as an application running on a host CPU.

In the case of a Wi-Fi or 802.1X network, there are numerous options for deploying the EAP client (used by a WPA supplicant) on the host. However, it should be kept in mind that the USIM or EIF doesn't only handle authentication but also stores the subscriber profile. Simply replacing the USIM with a WPA supplicant would therefore be insufficient in the case of cellular networks. The full functionality of the EIF is needed, specifically for providing access to the subscriber profile and terminating EAP sessions while being deployed as part of the OT domain. However, an existing WPA supplicant could be part of the EIF implementation and provide the required EAP client functionality.

The fact that the EIF forms part of the OT domain also means that methods defined and executed within the OT domain are used to provision the EIF in the device.

It should be noted that deploying the EIF outside an UICC in the operational domain doesn't necessarily lower the security bar. The EIF could integrate the industrial device's secure element (using, for instance, the Generic Trust Anchor API) to provide a high level of security.

4.3.4 NSSAA and Secondary Authentication

3GPP has also foreseen that 5G networks will be operated in environments with multiple stakeholders and that authentication and authorization decisions may not necessarily be made by a single entity in the network. This is the reason for introducing the concepts of network slice specific authentication and authorization (NSSAA) and secondary authentication (also known as data network (DN) authentication or protocol data unit (PDU) authentication). They are applicable to both SNPN and PNI-NPN deployment models.

In the case of NSSAA, in order for a device to access a certain logical partition of the 5G network (known as a slice) it may need to perform authentication and authorization via an additional authentication, authorization and accounting (AAA) server that could be outside of the 5G system and operated by the OT company. NSSAA doesn't replace primary authentication; it is optionally executed in addition to it.

Access to certain LAN or data center resources (which are grouped into a DN) also requires secondary authentication and authorization by the DN owner's AAA server. However, these don't replace primary authentication either, being optionally executed in addition to it instead. They use the EAP framework. Arbitrary EAP methods can be used between the device and a AAA server. 3GPP doesn't define any requirements for the EAP method or specify how identities and credentials for these new authentication types should be processed or handled on an industrial device. If these additional authentication methods are required, an authentication client needs to be deployed as part of the industrial device's OT domain.

Neither NSSAA nor secondary authentication is related to authentication or security procedures on the level of the industrial network protocol. Summing up, an industrial device could support up to four levels of authentication using different identities and credentials and key management approaches operated by different entities.

4.3.5 Summary

In the PNI-NPN model, if the OT takes advantage of user-plane services and slices provided by PLMN, for example, it's mandatory to use the USIM application for primary authentication on UICC or eUICC (including the iUICC form factor). This can't be avoided unless an agreement between the PLMN and the OT allows for an alternative mechanism.

Regarding the SNPN scenario, which is based on a standalone 5G network deployed and managed by OT, authentication for accessing the network may also use USIMs. Besides adopting the existing SIM ecosystem, vertical industries could benefit from replacing the USIM with an authentication client that supports EAP-based authentication for non-AKA credentials. These could be deployed as part of the OT domain of an industrial device, in other words independently of a UICC or eUICC (including the iUICC form factor).

Due to the USIM's strong legacy and major role in cellular networks, an approach combining storage functionality for subscriber profiles (traditionally provided by the USIM) with EAP client functionality to create a new function called EIF can only be implemented if there is strong support and the remaining architectural and technical issues are resolved.

5. Industrial 5G Device Physical Reference Architecture

In this section, we discuss various aspects of the physical architecture of industrial 5G devices.

First we address several defining aspects of this architecture:

- The need for explosion protection for devices in hazardous areas
- Options for implementing storage of credentials
- The ability to use either a chipset or a module
- The existing radio module form factor standards
- Selection of a standalone application processor or one that is integrated with the radio chipset
- Selection of an interface between the application processor and radio module

Then we present reference architecture diagrams for the device types introduced in section 3.1.

5.1 Explosion Protection for Devices in Hazardous Areas

5.1.1 Introduction

Flammable gases and vapors can occur in processing plants of the petroleum and chemical industries, among others. An area that has or may have such an explosive atmosphere is called a hazardous area. Special precautions must be taken when installing and operating devices in areas of this type to prevent them from causing fires or explosions.

5.1.2 Classification of Zones

The IEC 60079 [9] series of international standards establishes various requirements for the development, installation, operation, etc. of devices in hazardous areas. The requirements of most regional regulations on electrical devices, including the ATEX directives and the EN 60079 standards in Europe, are based on the IEC 60079 standards.

Hazardous areas with explosive atmospheres are assigned to three types of zones depending on how often explosive conditions occur and how long they last. IEC 60079 also stipulates the kinds of explosion protection that devices used in each zone must have (see 5.1.3) in order to minimize the risks.

- Zone 0: An explosive atmosphere is present continuously, for long periods, or frequently (for example, inside a tank of flammable liquid).
- Zone 1: An explosive atmosphere is likely to occur occasionally during normal operation (for example, around relief valves that release flammable gas during normal operation).
- Zone 2: An explosive atmosphere is unlikely to occur during normal operation, and if it does occur will quickly dissipate (for example, parts of a plant's premises to which flammable gas may occasionally drift).

All of the device types presented in section 3.1 can be installed or used in zone 1, zone 2, or non-hazardous areas depending on the use cases, configuration, and the plant's policy, while typically only sensors and actuators can be used in zone 0 areas.

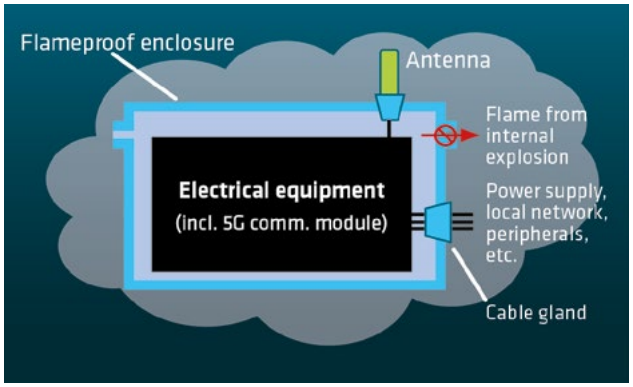
5.1.3 Types of Explosion Protection for Industrial Devices

This section introduces some of the explosion protection types defined by the IEC 60079 series of standards and describes the requirements that 5G communication modules would potentially have to meet for each level of explosion protection.

Protection by Flameproof Enclosure (Ex d)
An enclosure is considered to be flameproof (Ex d) if it is able to resist an internal explosion and prevent it from spreading to a surrounding explosive atmosphere. The requirements are

specified in IEC 60079-1. "Ex d" protection is usually provided for electrical equipment in zone 1 and 2 areas to prevent it from igniting an explosive atmosphere.

Figure 31: An example configuration of an industrial 5G device architecture protected by a flameproof enclosure

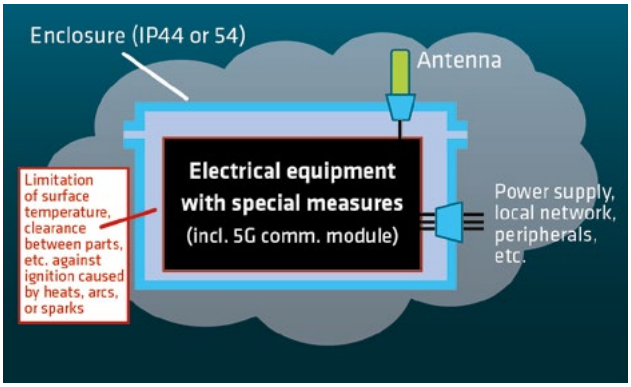


Internal electronic components, including 5G communication modules, could be vulnerable but a flameproof enclosure prevents the gas atmosphere surrounding them from igniting. If an internal explosion does occur, however, the electrical equipment inside the enclosure may be damaged by it.

Protection by Increased Safety (Ex e)

Increased safety or Ex e is an explosion protection concept that provides increased security against the risk of excessive temperatures and/or electrical arcs and sparks arising from electrical equipment in hazardous areas. IEC 60079-7 details the requirements for achieving this, such as impregnating coils, providing clearance between bare conductive parts, and so on. They make it possible to install and use equipment containing electronic circuits (like industrial 5G devices) under zone 2 conditions.

Figure 32: An industrial 5G device architecture with enhanced protection (Ex e)

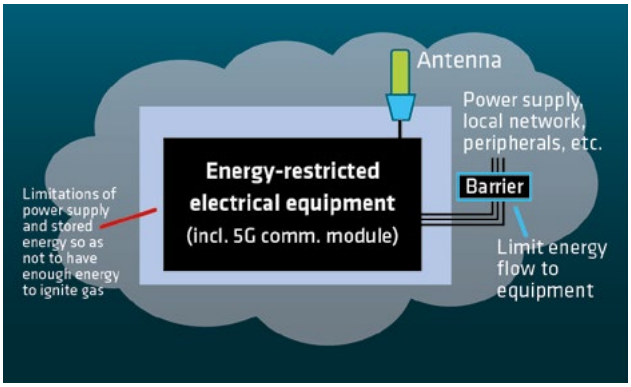


No surfaces of any internal parts, including 5G communication modules, should reach a temperature high enough to ignite an explosive atmosphere. IEC 60079 defines three groups of gases on the basis of their minimum ignition energies (IIA, IIB, and IIC) and six temperature classes based on the autoignition temperature of gases (T1 to T6), which must be taken into account when designing equipment for increased safety protection.

Protection by Intrinsic Safety (Ex i)

Protection by intrinsic safety or Ex i limits the electrical and thermal energy within equipment to a level below that at which ignition could be caused by sparking or heating, also under fault conditions. An apparatus called an intrinsic safety barrier limits the flow of energy supplied to the electrical equipment. The electrical equipment also restricts internal accumulation of energy. This protects areas with an explosive atmosphere and qualifies the electrical equipment as "intrinsically safe". The requirements are specified by IEC 60079-11. Equipment that qualifies for the highest level of protection defined by IEC 60079-11 ("Ex ia") may operate in zone 0 conditions.

Figure 33: An industrial 5G device architecture with protection by intrinsic safety (Ex i)



A 5G communication module designed for intrinsic safety must run on a limited energy supply. The energy stored in electronic circuits (like capacitors and inductors) of the equipment must also be limited. These constraints prevent the equipment from having enough energy to release an ignition spark in case a fault condition as defined by IEC 60079-11 occurs.

5.2 Physical Implementation for Storing Credentials

Management of credentials is an important aspect of industrial 5G network security. As shown in figure 34, credentials can be stored in different ways. It's also possible to combine several methods in the same industrial 5G device.

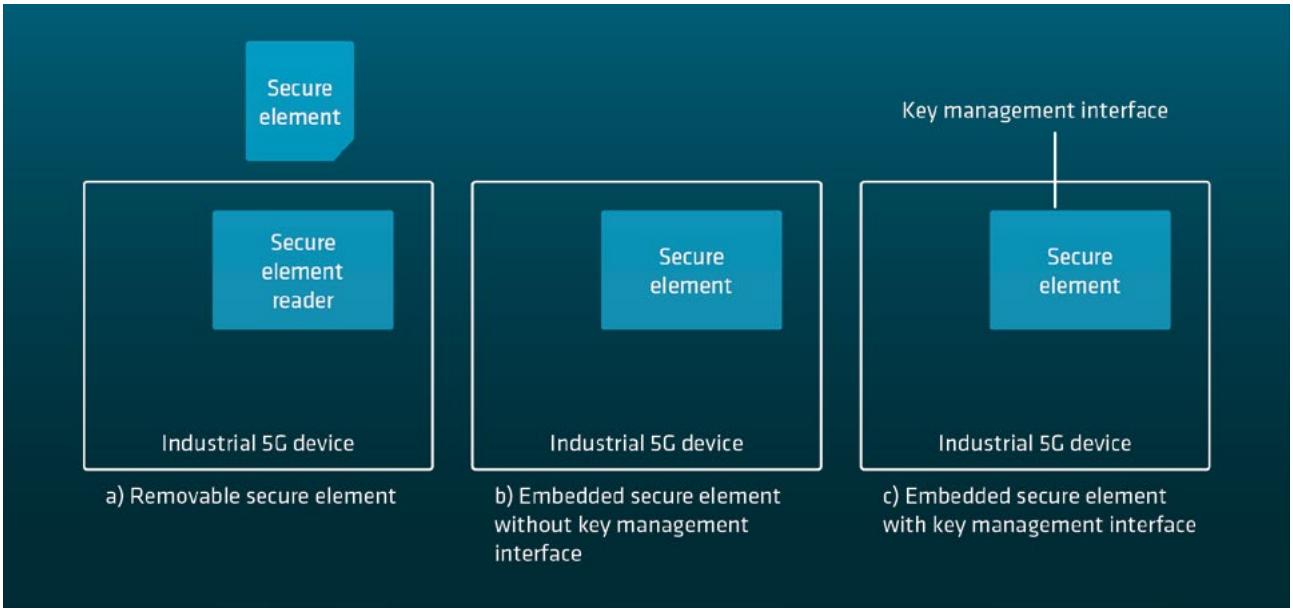
In this section, we describe the physical process of storing credentials in industrial 5G devices. How secure elements are connected inside an industrial 5G device is explained in chapter 4 from a logical perspective and further below in this section from a physical perspective.

To simplify the following description, here the concept of a "trust anchor" holding a device's initial credentials is introduced. It is also used to derive or securely download additional credentials.

It is possible to have two trust anchors, one for cellular authentication and another for application layer authentication. Alternatively, the same trust anchor can be used for both cellular and application layer authentication.

Please refer to section 4.3 for a full description of the various authentication methods supported by the 5G system.

Figure 34: Physical implementation options for storing credentials



5.2.1 Removable Secure Element

The trust anchor can be stored in a removable secure element. The secure element holding the trust anchor is then inserted into the industrial 5G device as shown in figure 34a.

A typical example of this is the UICC used to store the USIM application and possibly also other applications. The credentials are programmed into the UICC before the UICC is inserted into the device.

5.2.2 Embedded Secure Element Without Key Management Interface

It's also possible to integrate a secure element into an industrial 5G device. This is shown in figure 34b. In this case, the industrial 5G device lacks a key management interface. It is therefore necessary to program the trust anchor into the secure element before it is provided to the final customer.

A good example of this is the embedded or integrated UICC that supports GSMA's embedded SIM (eSIM) remote provisioning architecture.

5.2.3 Embedded Secure Element with Key Management Interface

Finally, it's possible to have an embedded secure element with a key management interface. This is shown in figure 34c. In this case, there is no need to load any credentials before supplying the industrial 5G device to the final customer.

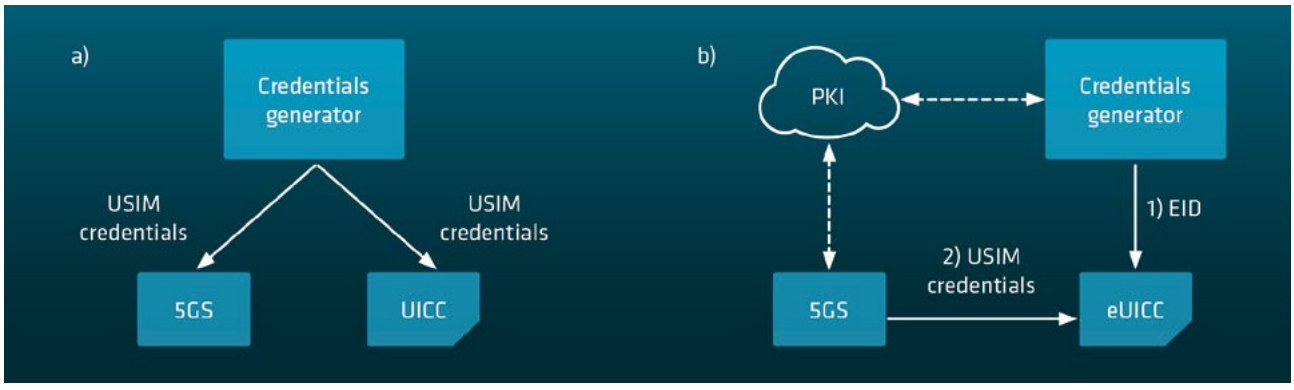
Depending on the provisioning protocol, the key management interface ensures the integrity and/or confidentiality of data arriving via the interface. This is commonly implemented as a local wired or short-range interface with optical, acoustic, near-field, or short-range wireless communication.

An example of this is a secure element used to store certificates for EAP authentication. The certificate is loaded using the simple certificate enrollment protocol (SCEP). Initially, a shared secret key is loaded via the key management interface. Then the certificate can be securely loaded via a wireless or wired interface into the secure element, also using SCEP.

5.2.4 Provisioning of Cellular Credentials

Provisioning of cellular credentials is generally accomplished in one of two different ways, as illustrated in figure 35.

Figure 35: Provisioning of cellular credentials



In figure 35a, USIM credentials are generated and then transferred to both an UICC and a 5GS. For public networks, UICCs are normally programmed at a central location and then physically transported to subscribers. For private networks, UICCs are normally programmed on site using a credentials generator.

Alternatively, credentials can be provisioned based on the GSMA's remote SIM provisioning framework. This is shown in figure 35b. First an embedded identity document (EID) is stored on the eSIM. The EID allows secure remote downloading of USIM credentials from a 5GS. The secure

remote download is based on a public key infrastructure to which all participating 5GS and credentials generators belong.

Both provisioning methods can be used for all three implementation options shown in figure 34

Finally, it's possible to use EAP-based cellular authentication. One example is EAP-TLS, which was introduced in 3GPP Release 15. EAP provides many different provisioning options, but credentials are usually provisioned via a key management interface on the device and/or using an automated credential management protocol.

5.3 Chipset Versus Module

Manufacturers have two main choices for implementing a 5G industrial device: a standard 5G modem chipset or a communication module containing one.

Choosing a 5G modem chipset makes it possible to develop a design that's optimized for a particular product. Fewer materials are also required, and there's no need to wait for modules to become available in the market before initiating product development. The downside is that it takes considerable expertise and experience to design and build a well-shielded and smoothly operating terminal (called UE in 3GPP terms). One critical aspect is radio frequency (RF) design, and another is meeting certification requirements. If the chipset is poorly designed, the 5G modem's performance, interoperability, and electromagnetic compatibility (EMC) will be compromised. This can result in unreliable connections, lower data throughput, increased latencies, and EMC certification challenges. Owing to these challenges, a chipset mainly makes sense for high-volume products.

5G communication modules are a recommended way of mastering these challenges. The module vendor takes care of RF calibration during production. The industrial 5G device manufacturer doesn't need to focus as much on RF design, since this is already largely covered by the module manufacturer. The interfaces provided by the module can simply be taken advantage of. What's more, it's possible to buy precertified 5G modules, thus greatly speeding up the certification process. In addition to these benefits, when integrating readily available 5G modules into a 5G industrial device it's possible to take advantage of ready-made modules that provide important processing resources (CPU, memory, I/Os) and can be used to implement core industrial device functionality beyond wireless cellular communications. For all of these reasons, the expectation is that the market will generally opt for the 5G module approach.

5.4 Radio Module Form Factor Standards

With regard to 5G module form factors and physical connections, two main categories of modules are available in the market: modules for soldering onto a printed circuit board (PCB) (like Land Grid Array (LGA) form factors) and pluggable modules (generally with an M.2 interface). The solderable modules typically include extra pins, making it possible to access more functionality of the 5G modem or use additional I/Os instead of a pluggable form factor with dedicated pins. On the other hand, no widely accepted specific form factor standard exists. This means that there is no guarantee that different 5G modules will be interchangeable. Pluggable form factors, with M.2 being a prominent format, have fewer pins but extensively standardized electrical properties. This lets 5G industrial device manufacturers upgrade their 5G industrial devices more easily later on without having to completely redesign them. Both form factor categories are technically feasible. At the end of the day, it's up to 5G industrial device manufacturers to decide which option meets their requirements better.

5.5 Standalone Versus Integrated Application Processor

Other architectural choices that 5G industrial device manufacturers must make include whether or not to integrate an application processor in the 5G module and if so, which other capabilities such as additional I/Os should be included. If the 5G module includes an adequately performing internal processor for executing customer-specific applications, its functionality can be expanded. Tasks normally assumed by dedicated external hardware, such as control applications (PLC, DCS, or motion or robot controllers), artificial intelligence algorithms, or visualization, can also be carried out by the internal application processor.

This won't necessarily impact its communication performance. Both architectural approaches – with or without a processor – involve tradeoffs. For example:

- Using a module with an integrated application processor may result in a smaller, more compact BOM and
- designing the PCB is simpler, but
- the 5G industrial device manufacturer will have fewer hardware resources available and more limited choices for integrating the OT application software in the 5G module.

Which approach is better depends on the specific use case.

Due to the mentioned advantages of using a 5G communication module, going forward it is expected to be the most common model. The architectural choices for industrial 5G devices shown in section 5.7 therefore assume that this approach is taken. For simplicity's sake, several of the components that a real industrial device would have are left out here (such as mechanical plugs, a power source, a housing and so on). Although the architectures shown in the following examples lack an integrated application processor, it could be feasibly be included in all of them.

5.6 Interface Between Application Processor and Radio Module

If a standalone application processor is chosen, there must be an interface to the radio module. This interface needs to support both data transfer and time synchronization.

5.6.1 Data Interface

The data interface can be implemented as multiple physical interfaces. Common options include:

- UART serial interface
- Universal serial bus (USB)
- Peripheral component interconnect express (PCIe)

UART-based serial interfaces used to be extensively used in modems. Due to their limited throughput, however, today they are mainly found in applications where this isn't an issue.

USB provides greater speed than UART-based serial interfaces. USB 3.1 can support up to 10 Gbit/s. The mobile broadband interface model (MBIM) interface was published by the USB Implementers Forums to enable broadband data connectivity via USB for cellular devices.

PCIe is an alternative to USB that makes it possible to scale up the speed even further. It also gives vendors greater flexibility for implementing higher-layer protocols.

The interfaces just described are primarily intended for configuration and data transfer. They are less suited for time synchronization between the application processor and radio module.

5.6.2 Time Synchronization Interface

A dedicated hardware interface is commonly used for time synchronization with GNSS receivers and other applications. Called 1pps or PPS, it generates a pulse that accurately repeats at regular time intervals. The timing information for each pulse arrives via a data interface. Consequently, there are actually two interfaces in play: a low-level interface that generates a pulse every second with microsecond accuracy without indicating which second it is and a high-level interface that indicates the second of the day.

A 5G system supports both global time domain and working clock domain synchronization. The use of a pulsed time reference signal together with higher-layer messaging via a digital interface is an effective way to synchronize an industrial 5G device with a TSN grand master or 5G system clock.

A similar interface can be used for time synchronization between a physical layer network interface and a radio module or application processor. Here the purpose is accurate synchronization of transmitted and received data frames.

5.7 Generic Block Diagrams for Industrial 5G Devices and Interface Options

It is useful to categorize the available architectures for industrial devices based on their use cases. This approach

lets them be grouped according to their characteristics (such as power, interfaces, processing capabilities, and so on) for defining the properties of 5G communication modules. The scheme presented in section 3.1 is the basis for doing this.

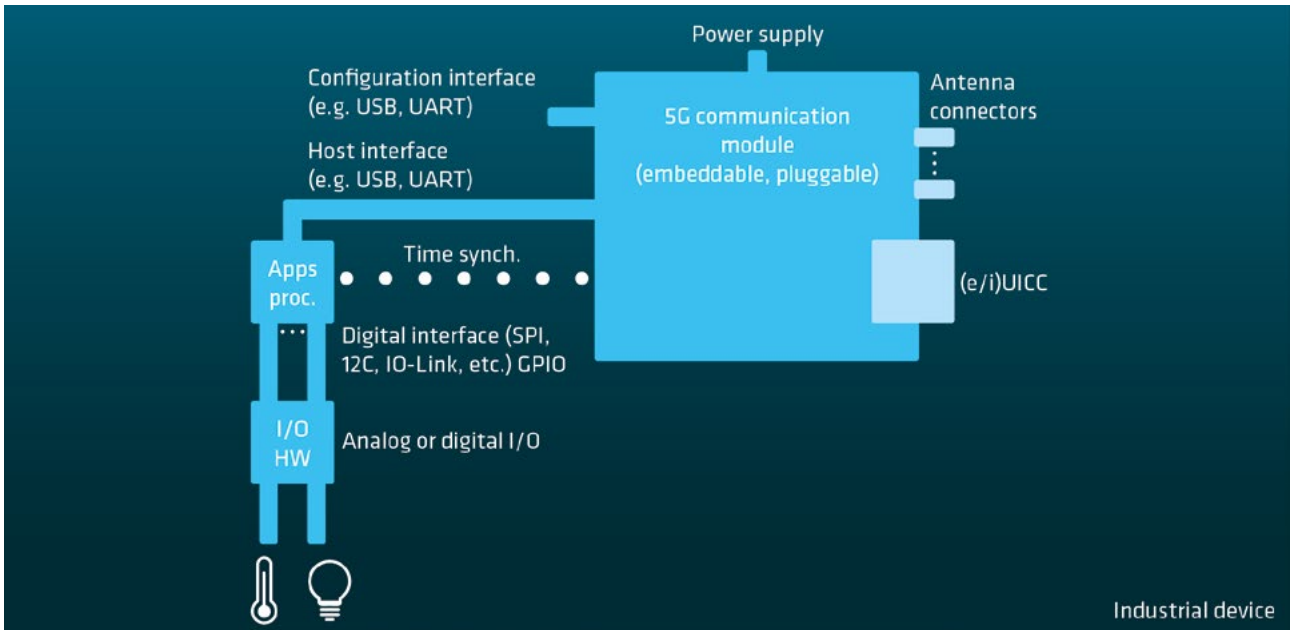
5.7.1 Industrial 5G Devices with Low-Power and Low-Latency Sensors/Actuators and 2D/3D Sensors

An industrial device of any of these types normally comprises a 5G communication module, interfaces for on-board or off-board connection of physical sensors/actuators or I/O transceivers, and optionally an application processor.

A 5G communication module typically has the following interfaces:

- Configuration interface (e.g. USB, UART)
- Host interface (e.g. USB, UART)
- Power supply
- Integrated antenna or antenna connector
- Time synchronization interface
- With an internal application processor, optionally digital interface(s) (SPI, I2C, UART etc.) for connecting I/O-protocol-specific transceiver(s) (such as IO-Link) or direct connection to analog or digital I/Os (such as GPIO, ADC, PWM) for directly connecting analog or digital sensors/actuators.
- Optional support for an (e/i)UICC/EAP identity function

Figure 36: Example 5G device architecture with sensor/actuator low-power industrial temperature sensor, 5G communication module, and external application processor

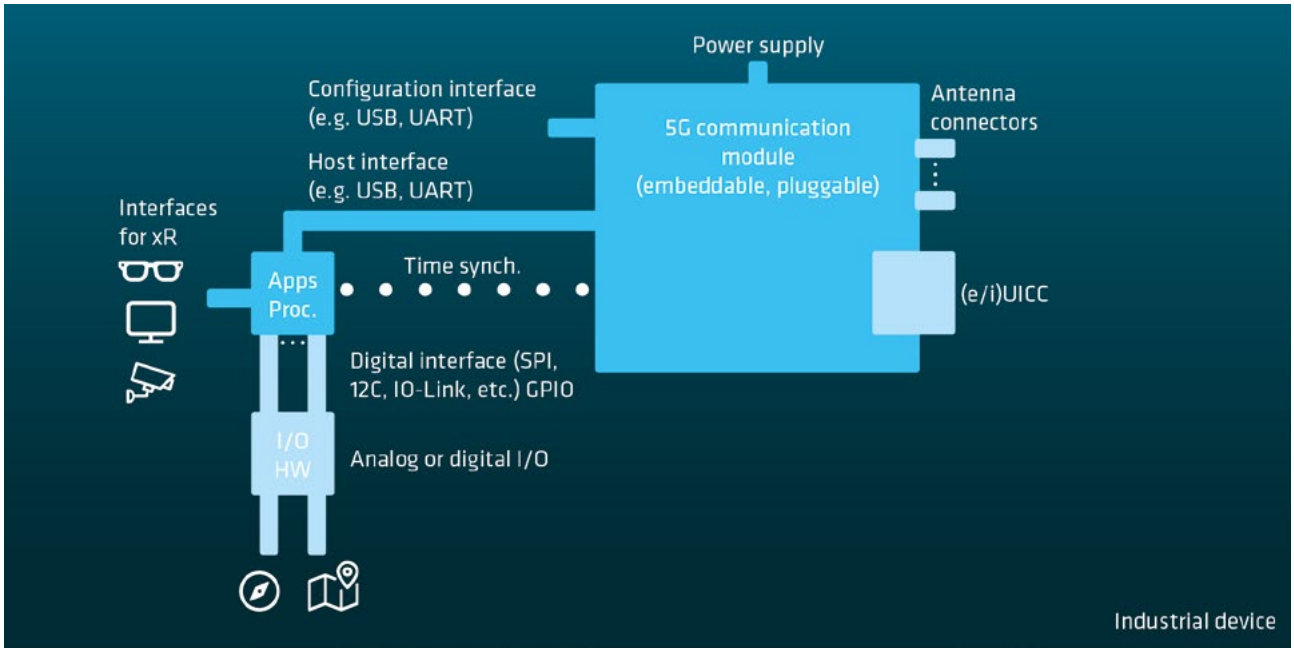


An architecture with integrated application processor and 5G communication module can accommodate devices with fewer and more compact components while minimizing power consumption. Conversely, an external processor solution permits partial reuse of existing software and layouts, thus potentially accelerating integration.

5.7.2 HMI and xR Devices

HMI and xR industrial devices typically contain – among other components – a 5G communication module and an internal application processor plus optional sensor/actuator hardware and integrated audio/visual components and/or interfaces to external A/V components. The processor (possibly including accelerators for xR processing) needs to be adequate for supporting visualization and xR applications.

Figure 37: HMI/xR industrial 5G device architecture incorporating a 5G communication module with external application processor



The 5G communication module in industrial devices of the HMI and xR category typically has the following interfaces:

- Configuration interface (e.g. USB, UART)
- Host interface (e.g. USB, UART)
- Power supply
- Integrated antenna or antenna connector
- Time synchronization interface
- Optionally, in case of module-internal applications processor: digital interface(s) (SPI, I2C, UART etc.) for connecting an I/O-protocol-specific transceiver (such as IO-Link) or directly linking to analog or digital I/Os (such as GPIO, ADC, PWM) for directly connecting analog or digital sensors/actuators.
- Optional support for an (e/i)UICC/EAP identity function

5.7.3 Gateways and PLCs/Controllers

5G industrial devices of these types typically comprise of – among other components – a 5G communication module plus I/O hardware for attaching the local fieldbus.

The 5G communication module has the following interfaces:

- Configuration interface (USB, UART)
- Host interface (USB, UART)
- Power supply
- Integrated antenna or an antenna connector
- Time synchronization interface
- Optionally, in the case of an internal application processor: a digital interface (such as PCIe, SPI, I2C, UART, etc.) for connecting to a local fieldbus (such as controllers or transceivers for Profibus, ModBus RTU, CAN bus, etc) and optionally GPIO for directly connecting digital sensors
- Optional support for an (e/i)UICC/EAP identity function

The application processor delivers enough performance to act as a proxy between the connected or preexisting fieldbuses and a higher-level control system (such as a PLC) that is reachable via the 5G system. In some categories, two architectural choices are possible: one with an integrated application processor and the other with an external application processor.

Figure 38 shows an example physical architecture for these device types.

5.7.4 TSN Port Industrial 5G Devices

The main role of devices that fall into the industrial 5G device category of TSN ports is to act as gateways between one or more local industrial Ethernet segment(s) and superordinate control system(s). They therefore typically comprise – among other components – a 5G communication module plus I/O hardware for connecting the local Ethernet segment.

The 5G communication module must support the following interfaces:

- Configuration/host interface (e.g. USB, UART)
- Power supply
- Ethernet port(s) (MAC, PHY, or logical via a PCIe attachment)
- Integrated antenna or antenna connector
- UICC (e/i) /EAP identity function support

Figure 38: Gateway and PLC/controller industrial 5G device architecture involving a 5G communication module with external applications processor

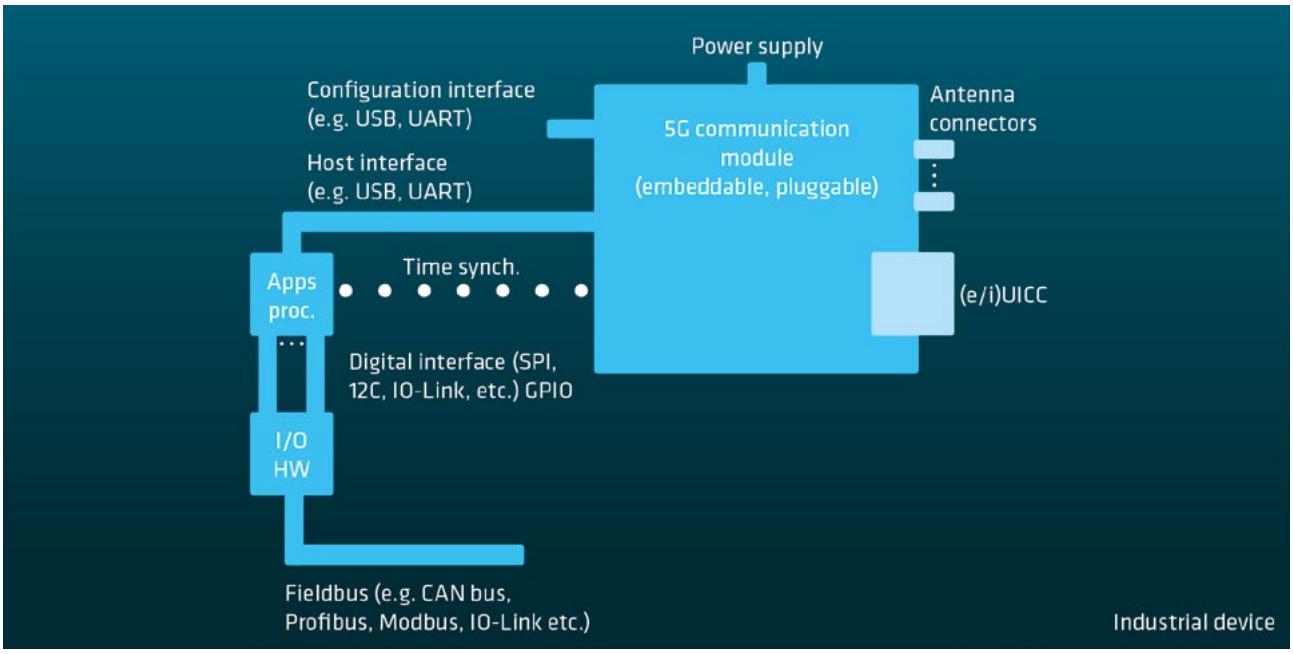
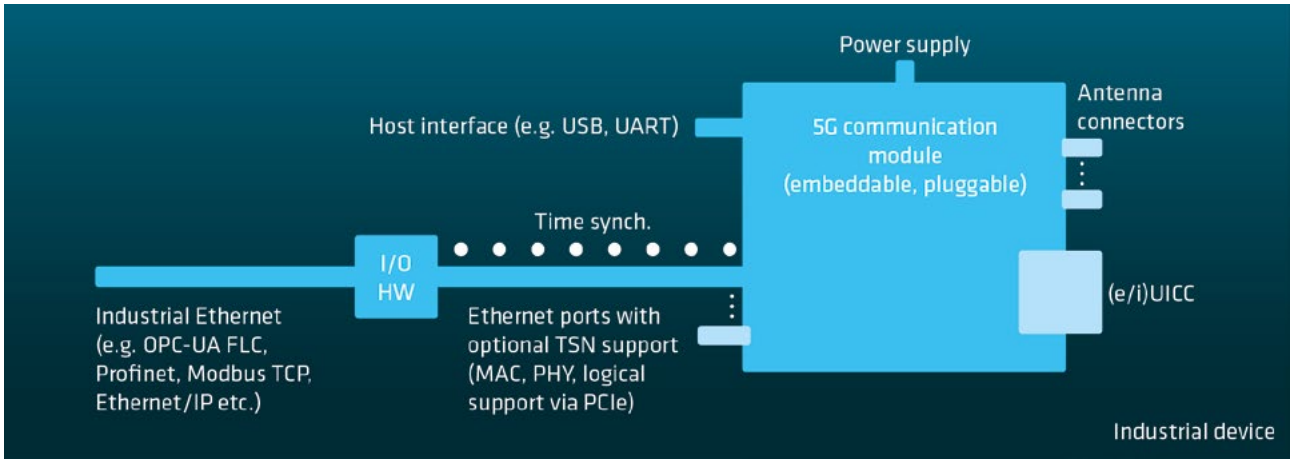


Figure 39: A TSN port industrial 5G device architecture with 5G communication module



6. Conclusions

In this white paper, we have discussed the architectural choices for industrial 5G devices. They depend on several criteria, including:

- The type of industrial 5G device
- Authentication based on USIM or EAP and how the credentials are stored in the industrial 5G device.
- The latency, throughput, and time synchronization needed for the industrial 5G device
- The environmental characteristics of the industrial 5G device. This includes protection from water, dust, vibration, and extreme temperatures as well as classification for operation in hazardous areas.
- Power characteristics and whether the industrial 5G device will be battery-powered or plugged into the grid

We have introduced a functional entity called an EAP identity function (EIF). The EIF holds the credentials needed for EAP authentication plus other relevant information that is otherwise stored in the USIM.

The GSMA has standardized remote SIM provisioning (RSP). From a technical perspective, the remote SIM provisioning standards may meet the need for provisioning operator credentials in a private network setting. However, the ecosystem for the remote SIM provisioning standards has been optimized for public operators.

The GSMA remote SIM provisioning (RSP) specifications let public operators remotely provision subscription profiles. It would be beneficial to adopt the remote SIM provisioning ecosystem to also meet the needs of private networks.

7. Definitions of Acronyms and Key Terms

3GPP
The 3rd Generation Partnership Project (3GPP) is an umbrella term for a consortium embracing a number of standards organizations worldwide that are collaborating to develop globally accepted specifications for mobile telecommunications. As its name implies, it was originally created to establish specifications for the third generation (3G) of mobile communication systems. It has continued working on subsequent generations, including the Fifth Generation (5G), which is considered in this white paper.

5G-ACIA
The 5G Alliance for Connected Industries and Automation is the globally leading organization for shaping and promoting industrial 5G.

5G network termination
Any 3GPP-defined device-side function involved in connecting to the 5G network and operating as part of the 5G system.

5G radio interface
A radio interface specified by 3GPP Release 15 or later, including 5GNR and E-UTRAN.

AAA
Authentication, authorization, and accounting.

DS-TT
Device-side TSN translator.

EAP
The extensible authentication protocol, defined in RFC 3748.

EAP-AKA
Extensible authentication protocol – authentication and key agreement, defined in RFC 4187.

EAP-AKA'
Extensible authentication protocol – authentication and key agreement, defined in RFC 5448.

EID
eUICC identifier. It uniquely identifies an eUICC and is cryptographically protected by a private key and a corresponding public key certificate issued within a public key infrastructure operated by the GSMA. During remote provisioning of SIM profiles to an eUICC, the EID is used to ensure that the profile is correctly deployed.

EIF
EAP identity function. A new functional element of the 5G security architecture, introduced in this white paper, that has not yet been standardized. The EIF is similar to a USIM, with the main differences being that it consists of EAP client functionality and there are no requirements for its deployment. In other words, the EIF can be deployed independently of a UICC and optimized to meet the needs of an industry vertical.

eSIM
Embedded-SIM. This conceptual term was introduced by the GSMA to describe the ability to provision a device with a USIM in electronic form by deploying a SIM profile (which can be received from a provisioning server) in a secure element called an eUICC. An eUICC can hold and execute multiple SIM profiles and switch between them. It is therefore the functional equivalent of several UICC cards, each of which has a USIM. The term eSIM is also often used to refer to an eUICC or, informally, to describe a UICC card with a soldered form factor, independently of its eUICC capabilities.[5]

eUICC
Embedded UICC. There are two competing definitions of this term:

(1) ETSI 103.465 defines it as an UICC that isn't readily accessible or replaceable, isn't intended to be removed or replaced in the terminal, and enables secure changing of subscriptions. Here the focus is on its permanent physical integration in a device.

(2) In the context of eSIM, a secure element's main characteristics are that it can be uniquely identified by an EID and provisioned with multiple SIM profiles (using remote SIM provisioning (RSP) capabilities defined by GSMA). An eUICC is thus the functional equivalent of multiple UICCs, each of which has a USIM. These characteristics are independent of

the form factor. It can even be deployed on a removable card (or on a soldered card or integrated into a system on a chip).

A UICC with a soldered form factor is sometimes also referred to as an eUICC, independently of whether it can change or provision subscriptions.

gPTP
Generalized precision time protocol. Defined in the IEEE 802.1AD series standards, it is related to PTP. One major difference is that gPTP only supports Ethernet transfer, while PTP also supports higher-layer protocols.

(g)PTP
Here this refers to either PTP, gPTP, or both.

GSMA
The GSM Association represents the interests of mobile network operators worldwide.

Industrial 5G device
An industrial device with a 3GPP-standardized 5G radio interface.

Industrial 5G device type
A type of industrial 5G device from an operational technology perspective.

iUICC
Integrated UICC. This term refers to a secure element that isn't deployed as a discretionary element but instead integrated into another element such as a CPU or a system on a chip. There is no consistent definition of the functional capabilities of an iUICC.

One possible interpretation of an iUICC is as a generic secure element able to host several secure applications, one of which could be a eUICC application. In this context, however, the more informal term iSSP (integrated smart secure platform) as standardized by ETSI SCP is replacing it.

Sometimes the term iUICC is informally used to refer to a eUICC with an integrated form factor and other times to refer to an UICC with an integrated form factor. To prevent misunderstandings, it should therefore only be used in a known and well-defined context.

Latency
The time it takes a message to travel from a sender to a receiver.

LGA
Land grid array: a type of surface-mount packaging for integrated circuits.

Local network termination
One or more functions for connecting a device to a local network.

Logical architecture
A structural design that includes as many details as possible without limiting the architecture to a particular technology or environment.

Non-real-time
A non-real-time system has functional requirements but no requirements to perform tasks within a specific period of time.

NW-TT
Network-side TSN translator.

OT
Operational technology: the technology needed to operate an industrial network. Common OT devices include sensors, actuators, and controllers.

Physical architecture
A structural design that provides enough detail to implement an architecture with a particular technology.

PLC
Programmable logical controller: an industrial controller used to control industrial machines or processes.

PTP
The precision time protocol defined by the IEEE 1588 series standards.

Real-time
A real-time system is characterized by the need to meet deadlines. In a hard real-time system, the data provided have no value if the deadline is exceeded. In a soft real-time system, data still have some value even if the deadline is exceeded.

Routing
When routing is used in the context of the Internet Protocol, it is the process of sending packets from a host on one network to a host on another network. When routing is used in the context of industrial protocols, it refers to the routing a higher-layer protocol across multiple industrial networks, possibly via different physical layers.

SCEP
Simple certificate enrollment protocol, defined in RFC 8894.

Secure element
A tamper-resistant dedicated platform consisting of hardware and software that is capable of securely hosting applications and their confidential and cryptographic data and providing a secure application execution environment [3]. Note that the secure element can have different form factors such as smart card, dedicated chip, or integrated in other components.

SIM
Subscriber identity module. An informal term that is typically used for a USIM deployed on a UICC. The term SIM card is ordinarily (but not exclusively) used if the UICC has the form factor of a removable card.

SIM profile
The entire content of a specific USIM in a serialized file format (including personalized data like SUPI or subscriber individual key). The SIM profile is used within the scope of remote SIM provisioning to provide the content of a USIM to a device and deploy the content to a eUICC. The format and content of SIM profiles are defined by the Trusted Connectivity Alliance.

Tagged data
Tagged data is data with relevant metadata. For example, a temperature could be expressed as a 16-bit integer. The metadata specify that the unit is Kelvin, Celsius, or Fahrenheit and indicate the offset and scaling factor, and may also include information on the source.

Time-aware
This term is used in IEEE Ethernet standards to describe a scheduler that considers absolute time (or the time of day) when scheduling frames in a queue. This is called a time-aware scheduler. The original Ethernet standards didn't consider absolute time for scheduling frames in queues, only relative time.

Time-sensitive
As used in the IEEE Ethernet standards, this attribute describes a network that supports real-time traffic. In this white paper we have used the term more broadly to also include communication in which timely delivery of information is important.

Time synchronization
Temporal synchronization of two or more clocks with one another.

Trust anchor
This holds a device's initial credentials and is also used to derive or securely download additional credentials.

TSN
Time-sensitive networking.

UICC
Universal integrated circuit card. Defined in 3GPP TS 31.101, it is used in USIM applications. It may exist with any of various form factors including removable or soldered cards, or integrated in another component such as a system on a chip (SoC) or CPU.

USIM
Universal subscriber identity module. A logical element of the 3GPP architecture that is defined by TS 31.102. It stores and provides access to all parameters comprising a subscriber profile. A USIM also provides security functions that are used by the mobile terminal for mutual authentication with the 5G network. 3GPP requires a USIM application to be deployed on every UICC.

8. References

[1] 5G-ACIA White Paper "Integration of 5G with Time-Sensitive Networking for Industrial Communications", published in February 2021.

[2] "Use Cases IEC/IEEE 60802", v1.3, https://www.ieee802.org/1/files/public/docs2018/60802-industrial-use-cases-0918-v13.pdf.

[3] ETSI TS 103 465 V15.0.0 (2019-05) "Smart Cards; Smart Secure Platform (SSP); Requirements Specification".

[4] 5G-ACIA White Paper "5G for Automation in Industry", published in July 2019.

[5] GSMA "eSIM White Paper: The what and how of remote SIM provisioning", published in March 2018.

[6] 3GPP Technical Specification 22.261, "Service requirements for the 5G system", v 16.14.1.

[7] Eike Lyczkowski, Andreas Wanjek, Christian Sauer, Wolfgang Kiess: "Wireless Communication in Industrial Applications", 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), September 2019, pp. 1392–1395.

[8] 3GPP Technical Specification 22.104, "Service requirements for cyber-physical control applications in vertical domains", v17.3.0, July 2020.

[9] International Electrotechnical Commission (IEC) standard "IEC 60079 Series Explosive Atmosphere Standards". Up to part 35.

5G-ACIA White Paper
Industrial 5G Devices – Architecture and Capabilities

Contact
5G Alliance for Connected Industries and
Automation (5G-ACIA), a Working Party of ZVEI
Lyoner Strasse 9
60528 Frankfurt am Main
Germany

Phone: +49 69 6302-424
Fax: +49 69 6302-319
Email: info@5g-acia.org
5g-acia.org

Published by
ZVEI e. V.
5G Alliance for Connected Industries and
Automation (5G-ACIA), a Working Party of ZVEI
zvei.org
5g-acia.org

March 2022

© ZVEI
This work, including all of its parts, is protected by copyright. Any use outside the strict limits of copyright law without the consent of the publisher is prohibited. This applies in particular to reproduction, translation, microfilming, storage, and processing in electronic systems. Although ZVEI has taken the greatest possible care in preparing this document, it accepts no liability for the content.

As of February 2022


Lenovo

LS telcom

MAVENIR

MC
TECHNOLOGIES

Microsoft

MITSUBISHI
ELECTRIC

MOXA

MUGLER
TELCO NETWORKS.

Nitto

NOKIA

NXP

orange

Panasonic

PEPPERL+FUCHS

PHOENIX
CONTACT

Qualcomm

RELIABLE
REAL-TIME
RADIO

Radisys

ROHDE&SCHWARZ

SAL
SILICON AUSTRIA LABS

salzburgresearch

Schneider
Electric

SICK

SIEMENS

SINTEF

SoftBank

SONY

life.augmented

Telefónica


TRUMPF

TZi

ublox

UPLINK

verizon

VIAVI
VIAVI Solutions

vodafone

VOLKSWAGEN
AKTIENGESELLSCHAFT

WAGO

Weidmüller

XITASO

YOKOGAWA

ZTE

5G-ACIA

5g-acia.org