

ETSI TS 122 104 V16.5.0 (2020-09)



5G;
Service requirements for cyber-physical control applications in
vertical domains
(3GPP TS 22.104 version 16.5.0 Release 16)



Reference

DTS/TSGS-0122104vg50

Keywords

5G

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommitteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2020.

All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members.

3GPP™ and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

oneM2M™ logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners.

GSM® and the GSM logo are trademarks registered and owned by the GSM Association.

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org/>).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Legal Notice

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under <http://webapp.etsi.org/key/queryform.asp>.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights	2
Legal Notice	2
Modal verbs terminology.....	2
Foreword.....	5
Introduction	5
1 Scope	6
2 References	6
3 Definitions, symbols and abbreviations	7
3.1 Definitions	7
3.2 Symbols.....	9
3.3 Abbreviations	9
4 Overview	9
4.1 Introduction	9
4.2 Activity patterns in automation	10
4.3 Communication attributes	10
4.4 Control systems and related communication patterns	10
4.5 Implications for 5G systems.....	11
5 Performance requirements.....	11
5.1 Overview	11
5.2 Periodic deterministic communication	12
5.3 Aperiodic deterministic communication	15
5.4 Non-deterministic communication	16
5.5 Mixed traffic.....	17
5.6 Clock synchronisation requirements	18
5.6.0 Description.....	18
5.6.1 Clock synchronisation service level requirements	18
5.6.2 Clock synchronisation service performance requirements.....	19
5.7 Positioning performance requirements.....	19
6 Ethernet applications	20
Annex A (informative): Summary of service performance requirements.....	22
A.1 About the vertical domains addressed in this Annex	22
A.2 Factories of the Future.....	22
A.2.1 Overview	22
A.2.2 Factory automation.....	23
A.2.2.1 Motion control	23
A.2.2.2 Control-to-control communication	25
A.2.2.3 Mobile robots.....	25
A.2.3 Process automation.....	27
A.2.3.1 Closed-loop control.....	27
A.2.3.2 Process monitoring	27
A.2.3.3 Plant asset management	28
A.2.4 Human machine interfaces	29
A.2.4.1 Mobile control panels	29
A.2.4.2 Augmented reality	31
A.2.5 Monitoring and maintenance.....	31
A.2.5.1 Remote access and maintenance.....	31
A.3 (void).....	33
A.4 Electric-power distribution.....	33

A.4.1	Overview	33
A.4.2	Primary frequency control	33
A.4.3	Distributed voltage control	34
A.4.4	Distributed automated switching for isolation and service restoration	34
A.4.5	Smart grid millisecond-level precise load control	36
A.5	Central power generation	37
A.5.1	Overview	37
A.5.2	Wind power plant network	37
Annex B (informative): Communication service errors		38
B.1	Introduction	38
B.2	Corruption	38
B.3	Unintended repetition	38
B.4	Incorrect sequence	38
B.5	Loss	39
B.6	Unacceptable deviation from target end-to-end latency	39
B.7	Masquerade	39
B.8	Insertion	39
B.9	Addressing	39
Annex C (informative): Characterising communication services		40
C.1	Modelling of communication in automation	40
C.1.1	Area of consideration	40
C.1.2	Logical link	41
C.1.2.1	Nature and function	41
C.1.2.2	Message transformation	42
C.1.2.3	Communication device	42
C.1.2.4	Communication system	42
C.2	Communication service description	43
C.2.1	Overview	43
C.2.2	Characteristic parameters	43
C.2.3	Influence quantities	45
C.3	Up time and up state vs. down state and down time	46
C.4	Timeliness as an attribute for timing accuracy	47
C.4.1	Overview	47
C.4.2	Network latency requirement formulated by use of timeliness	48
C.4.3	Timeliness	48
C.4.4	Deviation	48
C.4.5	Earliness	49
C.4.6	Lateness	50
C.4.7	Conclusion	50
C.5	Communication service terminology w.r.t. 5G network and vertical applications	50
Annex D (informative): 5G in industrial automation: different and multiple time domains for synchronization		53
D.1	Description	53
D.2	Merging of working clock domains	54
D.3	Time synchronization with 5G networks	55
Annex E (informative): Change history		56
History	57

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
 - 1 presented to TSG for information;
 - 2 presented to TSG for approval;
 - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

Introduction

The present document addresses a challenging class of vertical applications, namely cyber-physical control applications, which require very high levels of communication service availability, and some of them also require very low end-to-end latencies.

Real-time Ethernet is one of the established wireline communication technologies for cyber-physical control applications, and this specification identifies requirements that 5G systems must meet to support real-time Ethernet.

The present document provides new Stage 1 requirements based on the input from relevant stakeholders of the respective vertical domains.

1 Scope

The present document provides Stage 1 normative service requirements for 5G systems, in particular service requirements for cyber-physical control applications in vertical domains. In the context of the present document, cyber-physical systems are to be understood as systems that include engineered, interacting networks of physical and computational components; control applications are to be understood as applications that control physical processes.

Communication services supporting cyber-physical control applications need to be ultra-reliable and, in some cases, the end-to-end latency must be very low. Communication for cyber-physical control applications supports operation in various vertical domains, for instance industrial automation and energy automation.

The aspects addressed in the present document include:

- end-to-end service performance requirements and network performance requirements related to these end-to-end service performance requirements;
- support for LAN-type services specific to industrial/high performance use cases. Related Ethernet functionalities include, for example, those in IEEE 802.1Qbv.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [2] 3GPP TS 22.261: "Service requirements for the 5G system".
- [3] IEC 61784-3: "Industrial communication networks – profiles – part 3: functional fieldbuses – general rules and profile definitions".
- [4] BZKI, "Aspects of dependability assessment in ZDKI", June 2017.
- [5] BZKI, "Requirement Profiles in ZDKI", 2017.
- [6] IEC 61158: "Industrial communication networks – fieldbus specification", 2014.
- [7] IEC 61907, "Communication network dependability engineering".
- [8] Richard C. Dorf and Robert H. Bishop, "Modern Control Systems", Pearson, Harlow, 13th Edition, 2017.
- [9] Ernie Hayden, Michael Assante, and Tim Conway, "An Abbreviated History of Automation & Industrial Controls Systems and Cybersecurity", SANS Institute, <https://ics.sans.org/media/An-Abbreviated-History-of-Automation-and-ICS-Cybersecurity.pdf> {accessed: 2017-05-23}, 2014.
- [10] IEC 61512 "Batch control - Part 1: Models and terminology".
- [11] RESERVE project, Deliverable D1.3, ICT Requirements, <http://www.re-serve.eu/files/reserve/Content/Deliverables/D1.3.pdf>, September 2017.
- [12] RESERVE project, Deliverable D1.2, Energy System Requirements <http://www.re-serve.eu/files/reserve/Content/Deliverables/D1.2.pdf>, September 2017.

- [13] G. Garner, "Designing Last Mile Communications Infrastructures for Intelligent Utility Networks (Smart Grids)", IBM Australia Limited, 2010.
- [14] B. Al-Omar, B., A. R. Al-Ali, R. Ahmed, and T. Landolsi, "Role of Information and Communication Technologies in the Smart Grid", Journal of Emerging Trends in Computing and Information Sciences, Vol. 3, pp. 707-716, 2015.
- [15] H. Kagermann, W. Wahlster, and J. Helbig, "Recommendations for implementing the strategic initiative INDUSTRIE 4.0", Final report of the Industrie 4.0 working group, acatech – National Academy of Science and Engineering, Munich, April 2013.
- [16] IEC 62443-3-2: "Security for industrial automation and control systems - Part 3-2: Security risk assessment and system design", in progress.
- [17] IEC 62657-2: "Industrial communication networks - Wireless communication networks - Part 2: Coexistence management", 2017.
- [18] IEC 62657-1: "Industrial communication networks – Wireless communication networks – Part 1: Wireless communication requirements and spectrum considerations".
- [19] IEEE Std 802.1Q "Media Access Control (MAC) Bridges and Virtual Bridge Local Area Networks".

NOTE: IEEE Std 802.1Qbv-2015 "Enhancements for Scheduled Traffic" has been included into IEEE Std 802.1Q-2018.
- [20] IEEE, Use Cases IEC/IEEE 60802, 2018.
- [21] "IEEE Standard for Local and metropolitan area networks--Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks--Corrigendum 1: Technical and Editorial Corrections," IEEE Std 802.1AS-2011/Cor 1-2013 (Corrigendum to IEEE Std 802.1AS-2011), pp. 1-128, Sept 2013.
- [22] "IEEE Standard for Local and metropolitan area networks--Timing and Synchronization for Time-Sensitive Applications," IEEE Std 802.1AS-Rev/D7.3, pp. 1-502, August 2018.
- [23] 3GPP TS 22.289: "Mobile Communication System for Railways".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

characteristic parameter: numerical value that can be used for characterising the dynamic behaviour of communication functionality from an application point of view.

clock synchronicity: the maximum allowed time offset within a synchronisation domain between the sync master and any sync device .

NOTE 1: Clock synchronicity (or synchronicity) is used as KPI of clock synchronisation services.

NOTE 1A: Clock synchronicity is also referred to as clock (or time) synchronization precision.

clock synchronisation service: the service to align otherwise independent user-specific UE clocks.

communication service availability: percentage value of the amount of time the end-to-end communication service is delivered according to an agreed QoS, divided by the amount of time the system is expected to deliver the end-to-end service according to the specification in a specific area.

NOTE 2: The end point in "end-to-end" is assumed to be the communication service interface.

NOTE 3: The communication service is considered unavailable if it does not meet the pertinent QoS requirements. If availability is one of these requirements, the following rule applies: the system is considered unavailable if an expected message is not received within a specified time, which, at minimum, is the sum of maximum allowed end-to-end latency and survival time.

NOTE 4: This definition was taken from TS 22.261 [2].

communication service reliability: ability of the communication service to perform as required for a given time interval, under given conditions.

NOTE 5: Given conditions would include aspects that affect reliability, such as: mode of operation, stress levels, and environmental conditions.

NOTE 6: Reliability may be quantified using appropriate measures such as mean time between failures, or the probability of no failure within a specified period of time.

NOTE 7: This definition is based on IEC 61907 [7].

end-to-end latency: the time that takes to transfer a given piece of information from a source to a destination, measured at the communication interface, from the moment it is transmitted by the source to the moment it is successfully received at the destination.

NOTE 8: This definition was taken from TS 22.261 [2].

error: discrepancy between a computed, observed or measured value or condition and the true, specified or theoretically correct value or condition.

NOTE 9: taken from IEC 61784-3 [3].

factory automation: automation application in industrial automation branches typically with discrete characteristics of the application to be automated with specific requirements for determinism, low latency, reliability, redundancy, cyber security, and functional safety.

NOTE 10: Low latency typically means below 10 ms delivery time.

NOTE 11: This definition is taken from IEC 62657-1 [18].

global clock: a user-specific synchronization clock set to a reference timescale such as the International Atomic Time.

influence quantity: quantity not essential for the performance of an item but affecting its performance.

process automation: automation application in industrial automation branches typically with continuous characteristics of the application to be automated with specific requirements for determinism, reliability, redundancy, cyber security, and functional safety.

NOTE 12: This definition is taken from IEC 62657-1 [18].

service area: geographic region where a 3GPP communication service is accessible.

NOTE 13: The service area can be indoors.

NOTE 14: For some deployments, e.g., in process industry, the vertical dimension of the service area can be considerable.

NOTE 15: This definition was taken from TS 22.261 [2].

survival time: the time that an application consuming a communication service may continue without an anticipated message.

sync device: device that synchronizes itself to the master clock of the synchronization domain.

sync master: device serving as the master clock of the synchronization domain.

transfer interval: time difference between two consecutive transfers of application data from an application via the service interface to 3GPP system.

NOTE 16: This definition is based on subclause 3.1.85 in IEC 62657-2 [17].

user experienced data rate: the minimum data rate required to achieve a sufficient quality experience, with the exception of scenario for broadcast like services where the given value is the maximum that is needed.

NOTE 17: This definition was taken from TS 22.261 [2].

vertical domain: an industry or group of enterprises in which similar products or services are developed, produced, and provided.

working clock: a user-specific synchronization clock for a localized set of UEs collaborating on a specific task or work function.

3.2 Symbols

For the purposes of the present document, the following symbols apply:

<symbol> <Explanation>

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

CSIF	Communication Service Interface
EPON	Ethernet Passive Optical Network
FIFO	First In, First Out
GOOSE	Generic Object-Oriented Substation Event
HCL	Higher Communication Layer
HMI	Human Machine Interface
IMU	Inertial Measurement Unit
LCL	Lower Communication Layer
PMU	Phasor Measurement Unit

4 Overview

4.1 Introduction

For the purpose of this document, a vertical domain is a particular industry or group of enterprises in which similar products or services are developed, produced, and provided. Automation refers to the control of processes, devices, or systems in vertical domains by automatic means. The main control functions of automated control systems include taking measurements, comparing results, computing any detected or anticipated errors, and correcting the process to avoid future errors. These functions are performed by sensors, transmitters, controllers, and actuators.

In the context of this document, cyber-physical systems are referred to as systems that include engineered, interacting networks of physical and computational components. Cyber-physical control applications are to be understood as applications that control physical processes. Cyber-physical control applications in automation follow certain activity patterns, which are open-loop control, closed-loop control, sequence control, and batch control (see Clause 4.2).

Communication services supporting cyber-physical control applications need to be ultra-reliable, dependable with a high communication service availability, and often require low or (in some cases) very low end-to-end latency.

Communication in automation in vertical domains follows certain communication patterns. The most well-known is periodic deterministic communication, others are aperiodic deterministic communication and non-deterministic communication (see Clause 4.3).

Communication for cyber-physical control applications supports operation in various vertical domains, for instance industrial automation and energy automation. This document addresses service requirements for cyber-physical control applications and supporting communication services from the vertical domains of factories of the future (smart

manufacturing), electric power distribution, and central power generation. Service requirements for cyber-physical control applications and supporting communication services for rail-bound mass transit are addressed in TS 22.289 [23].

4.2 Activity patterns in automation

Open-loop control: The salient aspect of open-loop control is the lack of feedback from the output to the control; when providing commands to an actuator, it is assumed that the output of the influenced process is predetermined and within an acceptable range. This kind of control loop works if the influences of the environment on process and actuator are negligible. Also, this kind of control is applied in case unwanted output can be tolerated [8].

Closed-loop control: Closed-loop control enables the manipulation of processes even if the environment influences the process or the performance of the actuator changes over time. This type of control is realised by sensing the process output and by feeding these measurements back into a controller [8].

Sequence control: Sequence control may either step through a fixed sequence or employ logic that performs different actions based on various system states and system input [8]. Sequence control can be seen as an extension of both open-loop and closed-loop control, but instead of achieving only one output instance, an entire sequence of output instances can be produced [9].

Batch control: Batch processes lead to the production of finite quantities of material (batches) by subjecting input materials to a defined order of processing actions by use of one or more pieces of equipment [10].

4.3 Communication attributes

Communication in automation can be characterised by two main attributes: periodicity and determinism.

Periodicity means that a transmission interval is repeated. For example, a transmission occurs every 15 ms. Reasons for a periodical transmission can be the periodic update of a position or the repeated monitoring of a characteristic parameter. Most periodic intervals in communication for automation are rather short. The transmission is started once and continuous unless a stop command is provided.

An aperiodic transmission is, for example, a transmission which is triggered instantaneously by an event, i.e., events are the trigger of the transmission. Events are defined by the control system or by the user. Example events are:

- Process events: events that come from the process when thresholds are exceeded or fallen below, e.g., temperature, pressure, level, etc.
- Diagnostic events: events that indicate malfunctions of an automation device or module, e.g., power supply defective; short circuit; too high temperature; etc.
- Maintenance events: events based on information that indicates necessary maintenance work to prevent the failure of an automation device.

Most events, and especially alarms, are confirmed. In this context, alarms are messages that inform a controller or operator that an event has occurred, e.g., an equipment malfunction, process deviation, or other abnormal condition requiring a response. The receipt of the alarm is acknowledged usually within a short time period by the application that received the alarm. If no acknowledgment is received from the target application after a preset time, the so-called monitoring time, the alarm is sent again after a preset time or some failure response action is started.

Determinism refers to whether the delay between transmission of a message and receipt of the message at the destination address is stable (within bounds). Usually, communication is called deterministic if it is bounded by a given threshold for the latency/transmission time. In case of a periodic transmission, the variation of the interval is bounded.

4.4 Control systems and related communication patterns

There are preferences in the mapping between the type of control and the communication pattern. Open-loop control is characterised by one or many messages sent to an actuator. These can be sent in a periodic or an aperiodic pattern. However, the communication means used need to be deterministic since typically an activity response from the receiver and/or the receiving application is expected.

Closed-loop control produces both periodic and aperiodic communication patterns. Closed-loop control is often used for the control of continuous processes with tight time-control limits, e.g., the control of a printing press. In this case, one typically relies on periodic communication patterns. Note that in both the aperiodic and periodic case, the communication needs to be deterministic.

Logging of device states, measurements, etc. for maintenance purposes and such typically entails aperiodic communication patterns. In case the transmitted logging information can be time-stamped by the respective function, determinism is often not mandatory.

In practice, vertical communication networks serve a large number of applications exhibiting a wide range of communication requirements. In order to facilitate efficient modelling of the communication network during engineering and for reducing the complexity of network optimisation, traffic classes or communication patterns have been identified [6]. There are three typical traffic classes or communication patterns in industrial environments [6], i.e.,

- deterministic periodic communication: periodic communication with stringent requirements on timeliness of the transmission.
- deterministic aperiodic communication: communication without a preset sending time. Typical activity patterns for which this kind of communication is suitable are event-driven actions.
- non-deterministic communication: subsumes all other types of traffic, including periodic non-real time and aperiodic non-real time traffic. Periodicity is irrelevant in case the communication is not time-critical.

Some communication services exhibit traffic patterns that cannot be assigned to one of the above communication patterns exclusively (mixed traffic).

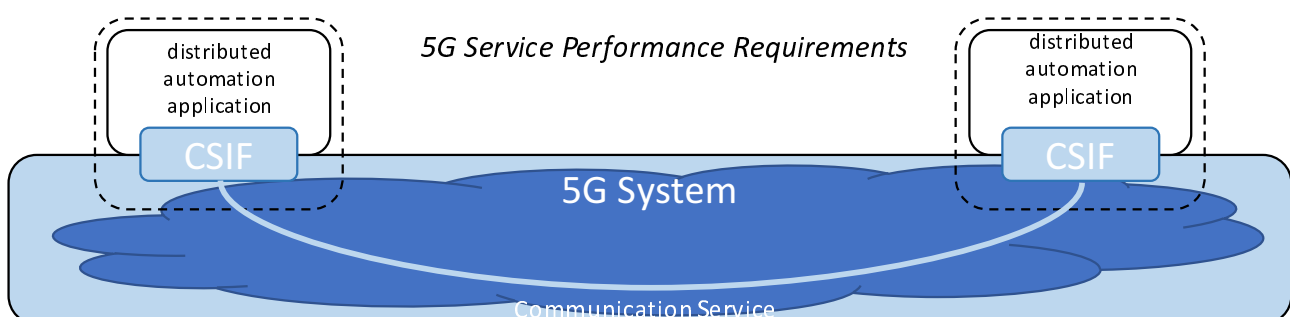
4.5 Implications for 5G systems

In order to be suitable for automation in vertical domains, 5G systems need to be dependable and flexible to meet specific KPIs to serve specific applications and use cases. They need to come with the system properties of reliability, availability, maintainability, safety, and integrity. What particular requirements each property needs to meet depends on the particularities of the domain and the use case. The requirements in this document provide various sets of performance criteria that need to be met to satisfactorily support different use cases of cyber-physical control applications used by various vertical markets.

5 Performance requirements

5.1 Overview

There are two fundamental perspectives concerning dependable communication in 5G systems: the end-to-end perspective of the communication services and the network perspective (see Figure 5.1-1).



CSIF – Communication Service Interface between distributed automation application/function and 5G system

Figure 5.1-1: Network perspective of 5G system

The Communication Service in Figure 5.1-1 may be implemented as a logical communication link between a UE on one side and a network server on the other side, or between a UE on one side and a UE on the other side.

In some cases, a local approach (e.g. network edge) is preferred for the communication service on the network side in order to reduce the latency, to increase communication service availability, or to keep sensitive data in a non-public network on the factory site.

The tables in Clauses 5.2 through 5.5 below provide sets of requirements, where periodicity and determinism are critical to meeting cyber-physical control application needs in various vertical scenarios. While many use cases have similar KPI values, the important distinction is that in order to meet the needs of different verticals and different uses, the 5G system will need to be sufficiently flexible to allow deployment configurations that can meet the different sets of KPIs specific to each use.

Communication service availability is considered an important service performance requirement for cyber-physical applications, especially for applications with deterministic traffic. The communication service availability depends on the latency and reliability (in the context of network layer packet transmissions, as defined in TS 22.261 [2]) of the logical communication link, as well as the survival time of the cyber-physical application (see Annex C.3 for further details on these relations).

The communication service reliability requirements also depend on the operation characteristics of the corresponding cyber-physical applications. Typically, the communication services critical for the automation application also come with stringent communication service reliability requirements. Note that the communication service reliability requirement has no direct relationship with the communication service availability requirement.

The "# of UEs" in the tables in clauses 5.2 to 5.5 is intended to give an indication of the UE density that would need to be served within a given service area (e.g. to understand the kind of capacity demand it puts on the 5G system).

Clock synchronisation is needed in many "vertical" use cases. The requirements and tables in Clause 5.6 provide specific criteria for managing time sensitive communications in an industrial environment.

High accuracy positioning is becoming essential for Factories of the Future. The reason for this is that tracking of mobile devices as well as mobile assets is becoming increasingly important in improving processes and increasing flexibility in industrial environments, Clause 5.7 provides positioning requirements for horizontal accuracy, availability, heading, latency and UE speed in an industrial use case scenario.

An example of the relationship between reliability (in the context of network layer packet transmissions, as defined in TS 22.261 [2]), survival time, and communication service availability of a logical communication link is illustrated in the following Table 5.1-1.

Table 5.1-1: Example of relationship between reliability (as defined in TS 22.261) and communication service availability when the survival time is equal to the transfer interval.

Communication service availability	Reliability (as defined in TS 22.261)
99,9999 %	99,9 %
99,999999 %	99,99 %
99,99999999 %	99,999 %
99,9999999999 %	99,9999 %
99,999999999999 %	99,99999 %

5.2 Periodic deterministic communication

Periodic deterministic communication is periodic with stringent requirements on timeliness and availability of the communication service. A transmission occurs every transfer interval. A description of periodic deterministic communication can be found in Clauses 4.3 and 4.4. Additional information on the underlying use cases of the sets of requirements in Table 5.2-1 can be found in Annex A. Further information on characteristic parameters and influence quantities used in Table 5.2-1 can be found in Annex C.

The 5G system shall be able to provide periodic deterministic communication with the service performance requirements for individual logical communication links that realise the communication services reported in Table 5.2-1.

Table 5.2-1: Periodic deterministic communication service performance requirements

Characteristic parameter				Influence quantity						Remarks
Communication service availability: target value (note 1)	Communication service reliability: mean time between failures	End-to-end latency: maximum (note 2)	Service bit rate: user experienced data rate	Message size [byte]	Transfer interval: target value	Survival time	UE speed	# of UEs	Service area (note 3)	
99,999 % to 99,99999 %	~ 10 years	< transfer interval value	–	50	500 µs	500 µs	≤ 75 km/h	≤ 20	50 m x 10 m x 10 m	Motion control (A.2.2.1)
99,9999 % to 99,999999 %	~ 10 years	< transfer interval value	–	40	1 ms	1 ms	≤ 75 km/h	≤ 50	50 m x 10 m x 10 m	Motion control (A.2.2.1)
99,9999 % to 99,999999 %	~ 10 years	< transfer interval value	–	20	2 ms	2 ms	≤ 75 km/h	≤ 100	50 m x 10 m x 10 m	Motion control (A.2.2.1)
99,9999 %	–	< 5 ms	1 kbit/s (steady state) 1,5 Mbit/s (fault case)	< 1500	< 60 s (steady state) ≥ 1 ms (fault case)	transfer interval	stationary	20	30 km x 20 km	Electrical Distribution – Distributed automated switching for isolation and service restoration (A.4.4); (note 5)
99,9999 % to 99,999999 %	~ 10 years	< transfer interval value		1 k	≤ 10 ms	10 ms	-	5 to 10	100 m x 30 m x 10 m	Control-to-control in motion control (A.2.2.2); (note 9)
> 99,9999 %	~ 10 years	< transfer interval value	–	40 to 250	1 ms to 50 ms (note 6) (note 7)	transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²	Mobile robots (A.2.2.3)
99,9999 % to 99,999999 %	~ 1 month	< transfer interval value	–	40 to 250	4 ms to 8 ms (note 7)	transfer interval value	< 8 km/h	TBD	50 m x 10 m x 4 m	Mobile control panels – remote control of e.g. assembly robots, milling machines (A.2.4.1); (note 9)
99,9999 % to 99,999999 %	~ 1 year	< transfer interval	–	40 to 250	< 12 ms (note 7)	12 ms	< 8 km/h	TBD	typically 40 m x 60 m; maximum 200 m x 300 m	Mobile control panels - remote control of e.g. mobile cranes, mobile pumps, fixed portal cranes (A.2.4.1); (note 9)
99,9999 % to 99,999999 %	≥ 1 year	< transfer interval value	–	20	≥ 10 ms (note 8)	0	typically stationary	typically 10 to 20	typically ≤ 100 m x 100 m x 50 m	Process automation – closed loop control (A.2.3.1)
99,999 %	TBD	~ 50 ms	–	~ 100	~ 50 ms	TBD	stationary	≤ 100 000	several km ² up to 100 000 km ²	Primary frequency control (A.4.2); (note 9)
99,999 %	TBD	~ 100 ms	–	~ 100	~ 200 ms	TBD	stationary	≤ 100 000	several km ² up to 100 000 km ²	Distributed Voltage Control (A.4.3) (note 9)

Characteristic parameter				Influence quantity						Remarks
Communication service availability: target value (note 1)	Communication service reliability: mean time between failures	End-to-end latency: maximum (note 2)	Service bit rate: user experienced data rate	Message size [byte]	Transfer interval: target value	Survival time	UE speed	# of UEs	Service area (note 3)	
> 99,9999 %	~ 1 year	< transfer interval value	–	15 k to 250 k	10 ms to 100 ms (note 7)	transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²	Mobile robots – video-operated remote control (A.2.2.3)
> 99,9999 %	~ 1 year	< transfer interval value	–	40 to 250	40 ms to 500 ms (note 7)	transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²	Mobile robots (A.2.2.3)
99,99 %	≥ 1 week	< transfer interval value	–	20 to 255	100 ms to 60 s (note 7)	≥ 3 x transfer interval value	typically stationary	≤ 10 000 to 100 000	≤ 10 km x 10 km x 50 m	Process monitoring (A.2.3.2), Plant asset management (A.2.3.3)

NOTE 1: One or more retransmissions of network layer packets may take place in order to satisfy the communication service availability requirement.
NOTE 2: Unless otherwise specified, all communication includes 1 wireless link (UE to network node or network node to UE) rather than two wireless links (UE to UE).
NOTE 3: Length x width (x height).
NOTE 4: (void)
NOTE 5: Communication includes two wireless links (UE to UE).
NOTE 6: This covers different transfer intervals for different similar use cases with target values of 1 ms, 1 ms to 10 ms, and 10 ms to 50 ms.
NOTE 7: The transfer interval deviates around its target value by < ± 25 %.
NOTE 8: The transfer interval deviates around its target value by < ± 5 %.
NOTE 9: Communication may include two wireless links (UE to UE)

5.3 Aperiodic deterministic communication

Aperiodic deterministic communication is without a pre-set sending time, but still with stringent requirements on timeliness and availability of the communication service. A description of aperiodic deterministic communication can be found in Clauses 4.3 and 4.4. Additional information on the underlying use cases of the sets of requirements in Table 5.3-1 can be found in Annex A. Further information on characteristic parameters and influence quantities used in Table 5.3-1 can be found in Annex C.

The 5G system shall be able to provide aperiodic deterministic communication with the service performance requirements for individual logical communication links that realise the communication services reported in Table 5.3-1.

Table 5.3-1: Aperiodic deterministic communication service performance requirements

Characteristic parameter (KPI)				Influence quantity					Remarks
Communication service availability	Communication service reliability: mean time between failures	Max Allowed End-to-end latency (note 1)	Service bit rate: user-experienced data rate	Message size [byte]	Survival time	UE speed	# of UEs	Service Area	
> 99,9999 %	~ 1 week	10 ms	> 10 Mbit/s			≤ 50 km/h	≤ 100	≤ 1 km ²	Mobile robots – video streaming (A.2.2.3)
99,9999 % to 99,999999 %	~ 1 month	< 30 ms	> 5 Mbit/s			< 8 km/h	TBD	TBD	Mobile control panels - parallel data transmission (A.2.4.1)
99,9999 %	–	< 50 ms	0,59 kbit/s 28 kbit/s	< 100	–	stationary	10~100 /km ²	TBD	Smart grid millisecond level precise load control (A.4.5)
> 99,9 %	~ 1 month	< 10 ms				< 8 km/h	≥ 3	20 m x 20 m x 4 m	Augmented reality; bi-directional transmission to image processing server (A.2.4.2)
NOTE 1: Unless otherwise specified, all communication includes 1 wireless link (UE to network node or network node to UE) rather than two wireless links (UE to UE).									
NOTE 2: (void)									

5.4 Non-deterministic communication

Non-deterministic communication subsumes all other traffic types than periodic/aperiodic deterministic communication. This includes periodic/aperiodic non-real-time traffic. A description of non-deterministic communication can be found in Clauses 4.3 and 4.4. Additional information on the underlying use cases of the sets of requirements in Table 5.4-1 can be found in Annex A. Further information on characteristic parameters and influence quantities used in Table 5.4-1 can be found in Annex C.

The 5G system shall be able to provide non-deterministic communication with the service performance requirements for individual logical communication links that realise the communication services reported in Table 5.4-1.

Table 5.4-1: Non-deterministic communication service performance requirements

Characteristic parameter (KPI)		Influence quantity			Remark
Communication service reliability: mean time between failures	Service bit rate: user-experienced data rate	UE speed	# of UEs	Service area (note)	
~ 1 month	≥ 1 Mbit/s	~ 0 km/h ≤ 75 km/h	≤ 100	50 m x 10 m x 10 m	Motion control - software updates (A.2.2.1)
	> 10 Mbit/s	≤ 50 km/h	≤ 100	≤ 1 km ²	Mobile robots; real-time video stream
NOTE: Length x width x height					

5.5 Mixed traffic

Mixed traffic cannot be assigned to one of the other communication patterns exclusively. Additional information on the underlying use cases of the sets of requirements in Table 5.5-1 can be found in Annex A. Further information on characteristic parameters and influence quantities used in Table 5.5-1 can be found in Annex C.

The 5G system shall be able to provide mixed traffic communication with the service performance requirements for individual logical communication links that realise the communication services reported in Table 5.5-1.

Table 5.5-1: Mixed traffic communication service performance requirements

Characteristic parameter (KPI)				Influence quantity			Remarks	
Communication service availability	Communication service reliability: mean time between failures	Max Allowed End-to-end latency (note 1)	Service bit rate: aggregate user-experienced data rate	Survival time	UE speed	# of UEs	Service Area	
99,9999999 %	~ 10 years	16 ms			stationary	< 1 000	several km ²	Wind power plant – control traffic (A.5.2);
NOTE 1: Unless otherwise specified, all communication includes 1 wireless link (UE to network node or network node to UE) rather than two wireless links (UE to UE).								
NOTE 2: (void)								

5.6 Clock synchronisation requirements

5.6.0 Description

Clock synchronicity, or time synchronization precision, is defined between a sync master and a sync device. The requirement on the synchronicity budget for the 5G system is the time error contribution between ingress and egress of the 5G system on the path of clock synchronization messages.

5.6.1 Clock synchronisation service level requirements

The 5G system shall support a mechanism to process and transmit IEEE1588v2 / Precision Time Protocol messages to support 3rd-party applications which use this protocol.

The 5G system shall support a mechanism to synchronise the user-specific time clock of UEs with a global clock.

The 5G system shall support a mechanism to synchronize the user-specific time clock of UEs with a working clock.

The 5G system shall support two types of synchronization clocks, the global time domain and the working clock domains.

The 5G system shall support networks with up to 32 working clock domains.

NOTE 1: The domain number (synchronization domain identifier) is defined with one octet in IEEE 802.1AS [22].

The 5G system shall support at least 2 simultaneous working clock domains on a UE.

The synchronicity budget for the 5G system within the global time domain shall not exceed 900 ns .

NOTE 2: The global time domain requires in general a precision of 1 μ s between the sync master and any device of the clock domain. Some use cases require only a precision of ≤ 100 μ s for the global time domain if a working clock domain with precision of ≤ 1 μ s is available.

NOTE 3: (void)

The synchronicity budget for the 5G system within a working clock domain shall not exceed 900 ns .

NOTE 4: The working clock domains require a precision of ≤ 1 μ s between the sync master and any device of the clock domain.

NOTE 5: Different working clock domains are independent and can have different precision.

The 5G system shall provide a media dependent interface for one or multiple 802.1AS sync domains [22].

The 5G system shall provide an interface to the 5G sync domain which can be used by applications to derive their working clock domain or global time domain (Reference Clock Model).

The 5G system shall provide an interface at the UE to determine and to configure the precision and time scale of the working clock domain.

5.6.2 Clock synchronisation service performance requirements

Table 5.6.2-1: Clock synchronization service performance requirements for the 5G System

User-specific clock synchronicity accuracy level	Number of devices in one Communication group for clock synchronisation	5GS synchronicity budget requirement (note)	Service area	Scenario
1	Up to 300 UEs	≤ 900 ns	≤ 100 m x 100 m	- Motion control - Control-to-control communication for industrial controller
2	Up to 10 UEs	< 10 μ s	≤ 2500 m ²	- High data rate video streaming
3	Up to 100 UEs	< 1 μ s	< 20 km ²	- Smart Grid: synchronicity between PMUs
NOTE: The clock synchronicity requirement refers to the clock synchronicity budget for the 5G system, as described in Clause 5.6.1.				

5.7 Positioning performance requirements

High accuracy positioning is becoming essential for Factories of the Future. The reason for this is that tracking of mobile devices as well as mobile assets is becoming increasingly important in improving processes and increasing flexibility in industrial environments.

Table 5.7-1 below lists typical scenarios and the corresponding high positioning requirements for horizontal accuracy, availability, heading, latency, and UE speed.

NOTE: The column on "Corresponding Positioning Service Level in TS 22.261" maps the scenarios listed in Table 5.7-1 to the service levels defined in TS 22.261 [2] which also include corresponding vertical accuracy requirements.

Table 5.7-1: Positioning performance requirements

Scenario	Horizontal accuracy	Availability	Heading	Latency for position estimation of UE	UE Speed	Corresponding Positioning Service Level in TS 22.261
Mobile control panels with safety functions (non-danger zones)	< 5 m	90 %	N/A	< 5 s	N/A	Service Level 2
Process automation – plant asset management	< 1 m	90 %	N/A	< 2 s	< 30 km/h	Service Level 3
Flexible, modular assembly area in smart factories (for tracking of tools at the work-place location)	< 1 m (relative positioning)	99 %	N/A	1 s	< 30 km/h	Service Level 3
Augmented reality in smart factories	< 1 m	99 %	< 0,17 rad	< 15 ms	< 10 km/h	Service Level 4
Mobile control panels with safety functions in smart factories (within factory danger zones)	< 1 m	99,9 %	< 0,54 rad	< 1 s	N/A	Service Level 4
Flexible, modular assembly area in smart factories (for autonomous vehicles, only for monitoring proposes)	< 50 cm	99 %	N/A	1 s	< 30 km/h	Service Level 5
Inbound logistics for manufacturing (for driving trajectories (if supported by further sensors like camera, GNSS, IMU) of indoor autonomous driving systems))	< 30 cm (if supported by further sensors like camera, GNSS, IMU)	99,9 %	N/A	10 ms	< 30 km/h	Service Level 6
Inbound logistics for manufacturing (for storage of goods)	< 20 cm	99 %	N/A	< 1 s	< 30 km/h	Service Level 7

6 Ethernet applications

6.1 Description

This section lists the requirements applicable to the 5G system for supporting cyber-physical applications using Ethernet.

For requirements pertaining to common, fundamental Ethernet transport requirements, and any requirements necessary to support the 5G LAN-type service, see Clause 6.24 in TS 22.261 [2].

6.2 Requirements

For infrastructure dedicated to high performance Ethernet applications, the 3GPP system shall support clock synchronisation defined by IEEE 802.1AS across 5G-based Ethernet links with PDU-session type Ethernet.

For infrastructure dedicated to high performance Ethernet applications, the 3GPP system shall support clock synchronisation defined by IEEE 802.1AS across 5G-based Ethernet links and other ethernet transports such as wired and optical (EPON.)

For infrastructure dedicated to high performance Ethernet applications, the accuracy of clock synchronisation should be better than 1µs.

For infrastructure dedicated to high performance Ethernet applications, the 3GPP system shall support enhancements for time-sensitive networking as defined by IEEE 802.1Q, e.g. time-aware scheduling with absolute cyclic time boundaries defined by IEEE 802.1Qbv [19], for 5G-based Ethernet links with PDU sessions type Ethernet.

For infrastructure dedicated to high performance Ethernet applications, absolute cyclic time boundaries shall be configurable for flows in DL direction and UL direction.

For infrastructure dedicated to high performance Ethernet applications, the 3GPP system shall support coexistence of hard-RT traffic following a time-aware schedule and lower priority traffic. The lower priority traffic cannot have a performance degrading impact on the hard-RT traffic.

The Ethernet transport service shall support routing based on information extracted from the Ethernet header information created based on 802.1Qbv.

Annex A (informative): Summary of service performance requirements

A.1 About the vertical domains addressed in this Annex

A vertical domain is an industry or group of enterprises in which similar products or services are developed, produced, and provided.

The vertical domains addressed in this Annex are

- Factories of the Future (A.2);
- electric-power distribution (A.4); and
- central power generation (A.5).

A.2 Factories of the Future

A.2.1 Overview

The manufacturing industry is currently subject to a fundamental change, which is often referred to as the "Fourth Industrial Revolution" or simply "Industry 4.0" [15]. The main goals of Industry 4.0 are—among others—the improvement of flexibility, versatility, resource efficiency, cost efficiency, worker support, and quality of industrial production and logistics. These improvements are important for addressing the needs of increasingly volatile and globalised markets. A major enabler for all this are cyber-physical production systems based on a ubiquitous and powerful connectivity, communication, and computing infrastructure. The infrastructure interconnects people, machines, products, and all kinds of other devices in a flexible, secure and consistent manner. Several different application areas can be distinguished:

- 1) **Factory automation:** Factory automation deals with the automated control, monitoring and optimisation of processes and workflows within a factory. This includes aspects like closed-loop control applications (e.g., based on programmable logic or motion controllers) and robotics, as well as aspects of computer-integrated manufacturing. Factory automation generally represents a key enabler for industrial mass production with high quality and cost-efficiency. Corresponding applications are often characterised by highest requirements on the underlying communication infrastructure, especially in terms of communication service availability, determinism, and latency. In the Factories of the Future, static sequential production systems will be more and more replaced by novel modular production systems offering a high flexibility and versatility. This involves many increasingly mobile production assets, for which powerful wireless communication and localisation services are required.
- 2) **Process automation:** Process automation refers to the control of production and handling of substances like chemicals, food & beverage, pulp, etc. Process automation improves the efficiency of production processes, energy consumption, and safety of the facilities. Sensors measuring process values, such as pressures or temperatures, are working in closed loops via centralised and decentralised controllers. In turn, the controllers interact with actuators, e.g., valves, pumps, heaters. Also, monitoring of attributes such as the filling levels of tanks, quality of material, or environmental data are important, as well as safety warnings or plant shut downs. Workers in the plant are supported by mobile devices. A process automation facility may range from a few 100 m² to several km², and the facility may be geographically distributed. Depending on the size, a production plant may have several 10 000 measurement points and actuators. Autarkic device power supply for years is needed in order to stay flexible and to keep the total costs of ownership low.
- 3) **HMIs and production IT:** Human-machine interfaces (HMIs) include all sorts of devices for the interaction between people and production facilities, such as panels attached to a machine or production line, but also standard IT devices, such as laptops, tablet PCs, smartphones, etc. In addition, augmented- and virtual-reality applications are expected to play an increasingly important role in future.

- 4) **Logistics and warehousing:** Organisation and control of the flow and storage of materials and goods in the context of industrial production. In this respect, intra-logistics is dealing with logistics within a certain property (e.g., within a factory), for example by ensuring the uninterrupted supply of raw materials on the shop floor level using automated guided vehicles (AGVs), fork lifts, etc. This is to be seen in contrast to logistics between different sites. Warehousing particularly refers to the storage of materials and goods, which is also getting more and more automated, for example based on conveyors, cranes and automated storage and retrieval systems.
- 5) **Monitoring and maintenance:** Monitoring of certain processes and/or assets in the context of industrial production without an immediate impact on the processes themselves (in contrast to a typical closed-loop control system in factory automation, for example). This particularly includes applications such as condition monitoring and predictive maintenance based on sensor data, but also big data analytics for optimising future parameter sets of a certain process, for instance. For these use cases, the data acquisition process is typically not latency-critical.

For each of these application areas, a multitude of potential use cases exists, some of which are outlined in the following subclauses. These use cases can be mapped to the given application areas (see Table A.2.1-1).

Table A.2.1-1: Mapping of the considered use cases (columns) to application areas (rows)

	Motion control	Control-to-control	Mobile control panels with safety	Mobile robots	Remote access and maintenance	Augmented reality	Closed-loop process control	Process monitoring	Plant asset management
Factory automation	X	X		X					
Process automation				X			X	X	X
HMIs and Production IT			X			X			
Logistics and warehousing		X		X					X
Monitoring and maintenance					X				

A.2.2 Factory automation

A.2.2.1 Motion control

A motion control system is responsible for controlling moving and/or rotating parts of machines in a well-defined manner, for example in printing machines, machine tools or packaging machines.

A schematic representation of a motion control system is depicted in Figure A.2.2.1-1. A motion controller periodically sends desired set points to one or several actuators (e.g., a linear actuator or a drive) which thereupon perform a corresponding action on one or several processes (in this case usually a movement or rotation of a certain component). At the same time, sensors determine the current state of the process(es), e.g. the current position and/or rotation of one or multiple components, and send the actual values back to the motion controller. This is done in a strictly cyclic and deterministic manner, such that during one application cycle the motion controller sends updated set points to all actuators, and all sensors send their actual values back to the motion controller. Nowadays, typically Industrial Ethernet technologies are used for motion control systems.

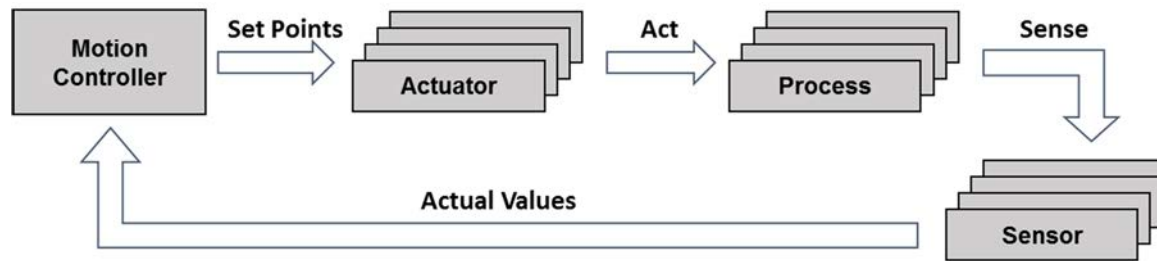


Figure A.2.2.1-1: Schematic representation of a motion control system

Table A.2.2.1-1: Service performance requirements for motion control

Use case #	Characteristic parameter				Influence quantity						
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: lower bound	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area (note)
1	99,999 to 99,99999	~ 10 years	< transfer interval value	–	50	500 μs – 500 ns	500 μs + 500 ns	500 μs	≤ 72 km/h	≤ 20	50 m x 10 m x 10 m
2	99,9999 to 99,999999	~ 10 years	< transfer interval value	–	40	1 ms – 500 ns	1 ms + 500 ns	1 ms	≤ 72 km/h	≤ 50	50 m x 10 m x 10 m
3	99,9999 to 99,999999	~ 10 years	< transfer interval value	–	20	2 ms – 500 ns	2 ms + 500 ns	2 ms	≤ 72 km/h	≤ 100	50 m x 10 m x 10 m

NOTE: Length x width x height.

Use cases one to three

Characteristic parameters and influence quantities for a communication service supporting the cyclic interaction described above.

A.2.2.2 Control-to-control communication

Control-to-control communication, i.e., the communication between different industrial controllers is already used today for different use cases, such as:

- large machines (e.g., newspaper printing machines), where several controls are used to cluster machine functions, which need to communicate with each other; these controls typically need to be synchronised and exchange real-time data;
- individual machines that are used for fulfilling a common task (e.g., machines in an assembly line) often need to communicate, for example for controlling and coordinating the handover of work pieces from one machine to another.

Typically, a control-to-control network has no fixed configuration of certain controls that need to be present. The control nodes present in the network often vary with the status of machines and the manufacturing plant. Therefore, hot-plugging support for different control nodes is important and often used.

Table A.2.2.2-1: Service performance requirements for control-to control communication in motion control

Use case #	Characteristic parameter			Influence quantity					
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval	Survival time	UE speed	# of UEs	Service area (note 1)
1 (note 2)	99,9999 to 99,999999	~ 10 years	< transfer interval value	1 k	≤ 10 ms	10 ms	stationary	5 to 10	100 m x 30 m x 10 m
NOTE 1: Length x width x height.									
NOTE 2: Communication may include two wireless links (UE to UE)									

Use case one

Control-to-control communication between different motion (control) subsystems, as addressed in Subclause A.2.2.1. An exemplary application for this are large printing machines, where it is not possible or desired to control all actuators and sensors by one motion controller only.

A.2.2.3 Mobile robots

Mobile robots and mobile platforms, such as automated guided vehicles, have numerous applications in industrial and intra-logistics environments and will play an increasingly important role in the Factory of the Future. A mobile robot essentially is a programmable machine able to execute multiple operations, following programmed paths to fulfil a large variety of tasks. This means, a mobile robot can perform activities like assistance in work steps and transport of goods, materials and other objects. Mobile robot systems are characterised by a maximum flexibility in mobility relative to the environment, with a certain level of autonomy and perception ability, i.e., they can sense and react with their environment.

Autonomous guided vehicles are a sub-group of mobile robots. These vehicles are driverless and used for moving materials efficiently within a facility. A detailed overview of the state of the art of autonomous-guided-vehicle systems is provided elsewhere in the literature [16].

Mobile robots are monitored and controlled from a guidance control system. Radio-controlled guidance control is necessary to get up-to-date process information, to avoid collisions between mobile robots, to assign driving jobs to the mobile robots, and to manage the traffic of mobile robots. The mobile robots are track-guided by the infrastructure with markers or wires in the floor or guided by own surround sensors, like cameras and laser scanners.

Mobile robot systems can be divided in operation in indoor, outdoor and both indoor and outdoor areas. These environmental conditions have an impact on the requirements of the communication system, e.g., the handover process, to guarantee the required cycle times.

Table A.2.2.3-1: Service performance requirements for mobile robots

Use case #	Characteristic parameter				Influence quantity							
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: lower bound	Transfer interval: target value (note)	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area
1	> 99,9999	~ 10 years	< target transfer interval value	–	40 to 250	– < 25 % of target transfer interval value	1 ms to 50 ms	+ < 25 % of target transfer interval value	target transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²
2	> 99,9999	~ 1 year	< target transfer interval value	–	15 k to 250 k	– < 25 % of target transfer interval value	10 ms to 100 ms	+ < 25 % of target transfer interval value	target transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²
3	> 99,9999	~ 1 year	< target transfer interval value	–	40 to 250	– < 25 % of target transfer interval value	40 ms to 500 ms	+ < 25 % of target transfer interval value	target transfer interval value	≤ 50 km/h	≤ 100	≤ 1 km ²
4	> 99,9999	~ 1 week	10 ms	> 10 Mbit/s	–	–	–	–	–	≤ 50 km/h	≤ 100	≤ 1 km ²

NOTE: The transfer interval is not so strictly periodic in these use cases. The transfer interval deviates around its target value within bounds. The mean of the transfer interval is close to the target value.

Use case one

Periodic communication for the support of precise cooperative robotic motion control (transfer interval: 1 ms), machine control (transfer interval: 1 ms to 10 ms), co-operative driving (10 ms to 50 ms).

Use case two

Periodic communication for video-operated remote control.

Use case three

Periodic communication for standard mobile robot operation and traffic management.

Use case four

Real-time streaming data transmission (video data) from a mobile robot to the guidance control system.

A.2.3 Process automation

A.2.3.1 Closed-loop control

In the closed-loop control use case for process automation, several sensors are installed in a plant and each sensor performs continuous measurements. The measurement data are transported to a controller, which takes decision to set actuators. The latency and determinism in this use case are crucial. This use case has very stringent requirements in terms of latency and service availability. The required service area is usually bigger than for motion control use cases. Interaction with the public network (e.g., service continuity, roaming) is not required.

Table A.2.3.1-1: Service performance requirements for closed-loop control in process automation

Use case #	Characteristic parameter			Influence quantity							
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval: lower bound	Transfer interval: target value	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area (note)
1	99,9999 to 99,999999	≥ 1 year	< target transfer interval value	20	– 5 % of target transfer interval value	≥ 10 ms	+ 5 % of target transfer interval value	0	typically stationary	typically 10 to 20	typically ≤ 100 m x 100 m x 50 m
NOTE: Length x width x height.											

Use case one

Several sensors are installed in a plant and each sensor performs continuous measurements. The measurement data are transported to a controller, which takes decision to set actuators.

A.2.3.2 Process monitoring

For process monitoring in the area of process automation, several sensors are installed in the plant to give insight into process or environmental conditions or inventory of material. The data are transported to displays for observation and/or to databases for registration and trending. The operation for this use case can be in a wide service area, and interaction with the public network (e.g., service continuity, roaming) may be required.

Table A.2.3.2-1: Service performance requirements for process monitoring

Use case #	Characteristic parameter			Influence quantity							
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval: lower bound	Transfer interval: target value (note 2)	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area (note 1)
1	99,99	≥ 1 week	< target transfer interval value	20	– 25 % of target transfer interval value	100 ms to 60 s	+ 25 % of target transfer interval value	matter of convenience; typically ≥ 3 x target transfer interval value	typically stationary	up to 10 000	typically ≤ 10 km x 10 km x 50 m
NOTE 1: Length x width x height.											
NOTE 2: The transfer interval is not so strictly periodic in these use cases. The transfer interval deviates around its target value within bounds. The mean of the transfer interval is close to the target value.											

Use case one

Several sensors are installed in the plant for providing insight into process or environmental conditions, or into inventory of material. The messages are transported to displays for observation and/or to databases for registration and trending.

A.2.3.3 Plant asset management

To keep a plant running, it is essential that the assets, such as pumps, valves, heaters, instruments, etc., are maintained. Timely recognition of any degradation and continuous self-diagnosis of components are used to support and plan maintenance. Remote software updates enhance and adapt the components to changing conditions and advances in

technology. The positioning requirements for this use case are mainly based on the typical scenarios where IoT devices (e.g., sensors) are giving insight into process or environmental conditions or inventory of material, asset management or maintenance. The operation for this use case can be in a wide service area, and interaction with the public network (e.g., service continuity, roaming) may be required.

Table A.2.3.3-1: Service performance requirements for plant asset management

Use case #	Characteristic parameter			Influence quantity					
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval	Survival time	UE speed	# of UEs	Service area (note)
1	99,99	TBD	< transfer interval value	20 to 255	several seconds	matter of convenience; typically $\geq 3 \times$ transfer interval value	typically stationary	$\leq 100\ 000$	typically $\leq 10\ \text{km} \times 10\ \text{km} \times 50\ \text{m}$
NOTE: Length x width x height.									

Use case one

To keep a plant running, it is essential that the assets, such as pumps, valves, heaters, and instruments are maintained. Timely recognition of any degradation and continuous self-diagnosis of components are used to support and plan maintenance. Remote software updates enhance and adapt the components to changing conditions and advances in technology.

A.2.4 Human machine interfaces

A.2.4.1 Mobile control panels

Control panels are crucial devices for the interaction between people and production machinery as well as for the interaction with moving devices. These panels are mainly used for configuring, monitoring, debugging, controlling and maintaining machines, robots, cranes or complete production lines. In addition to that, (safety) control panels are typically equipped with an emergency stop button and an enabling device, which an operator can use in case of a safety event in order to avoid damage to humans or machinery. When the emergency stop button is pushed, the controlled equipment immediately comes to a safe stationary position. Likewise, if a machine, robot, etc. is operated in the so-called special ‘enabling device mode’, the operator manually keeps the enabling device switch in a special stationary position. If the operator pushes this switch too much or releases it, the controlled equipment immediately comes to a safe stationary position as well. This way, it can be ensured that the hand(s) of the operator are on the panel (and not under a moulding press, for example), and that the operator does—for instance—not suffer from any electric shock or the like. A common use case for this ‘enabling device mode’ is the installation, testing or maintenance of a machine, during which other safety mechanisms (such as a safety fence) are deactivated.

Due to the criticality of these safety functions, safety control panels currently have mostly a wire-bound connection to the equipment they control. In consequence, there tend to be many such panels for the many machines and production units that typically can be found in a factory. With an ultra-reliable low-latency wireless link, it would be possible to connect such mobile control panels with safety functions wirelessly. This would lead to a higher usability and would allow for the flexible and easy re-use of panels for controlling different machines.

The cycle times of the control application depends on the process/machinery/equipment whose safety has to be ensured. For a fast-moving robot, for example, end-to-end latencies are lower than for slowly moving linear actuators.

Table A.2.4.1-1: Service performance requirements for mobile control panels

Use case #	Characteristic parameter				Influence quantity							
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: lower bound	Transfer interval: target value	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area (note)
1 (note 3)	99,9999 to 99,999999	~ 1 month	< target transfer interval value	–	40 to 250	– < 25 % of target transfer interval value	4 ms to 8 ms	+ 25 % of target transfer interval value	target transfer interval value	< 7,2 km/h	TBD	50 m x 10 m x 4 m
2 (note 3)	99,9999 to 99,999999	~ 1 month	< target transfer interval value	> 5 Mbit/s	–	– < 25 % of target transfer interval value	< 30 ms	+ 25 % of target transfer interval value	TBD	< 7,2 km/h	TBD	TBD
3 (note 3)	99,9999 to 99,999999	~ 1 year	< target transfer interval	–	40 to 250	– < 25 % of target transfer interval value	< 12 ms	+ 25 % of target transfer interval value	12 ms	< 7,2 km/h	TBD	typically 40 m x 60 m; maximum 200 m x 300 m

NOTE 1: Length x width (x height).
NOTE 2: The transfer interval is not so strictly periodic in these use cases. The transfer interval deviates around its target value within bounds. The mean of the transfer interval is close to the target value.
NOTE 3: Communication may include two wireless links (UE to UE)

Use case one

Periodic, bi-directional communication for remote control. Examples for controlled units: assembly robots; milling machines.

Use case two

Aperiodic data transmission in parallel to remote control (use case one).

Use case three

Periodic, bi-directional communication for remote control. Examples for controlled units: mobile cranes, mobile pumps, fixed portal cranes.

A.2.4.2 Augmented reality

It is envisioned that in future smart factories and production facilities, people will continue to play an important and substantial role. However, due to the envisaged high flexibility and versatility of the Factories of the Future, shop floor workers should be optimally supported in getting quickly prepared for new tasks and activities and in ensuring smooth operations in an efficient and ergonomic manner. To this end, augmented reality may play a crucial role, for example for the following applications:

- monitoring of processes and production flows;
- step-by-step instructions for specific tasks, for example in manual assembly workplaces;
- ad hoc support from a remote expert, for example for maintenance or service tasks.

In this respect, especially head-mounted augmented-reality devices with see-through display are very attractive since they allow for a maximum degree of ergonomics, flexibility and mobility and leave the hands of workers free for other tasks. However, if such augmented-reality devices are worn for a longer period (e.g., one work shift), these devices have to be lightweight and highly energy-efficient while at the same time they should not become very warm. A very promising approach is to offload complex (e.g., video) processing tasks to the network (e.g., an edge cloud) and to reduce the augmented-reality head-mounted device's functionality. This has the additional benefit that the augmented-reality application may have easy access to different context information (e.g., information about the environment, production machinery, the current link state, etc.) if executed in the network.

Table A.2.4.2-1: Service performance requirements for augmented reality in human-machine interfaces

Use case #	Characteristic parameter			Influence quantity	
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	UE speed	Service area (note)
1	> 99,9	~ 1 month	< 10 ms	< 8 km/h	20 m x 20 m x 4 m
NOTE: Length x width x height.					

Use case one

Bi-directional message transmission between an augmented-reality device and an image processing server.

A.2.5 Monitoring and maintenance

A.2.5.1 Remote access and maintenance

In factories of the future, there are needs to perform remote access and maintenance to devices and entities, for instance, by remote control centres. The devices and entities might be installed at geographically distributed locations. These devices typically have firmware/software which needs to be updated occasionally. Maintenance information also needs to be collected and distributed from/to these devices periodically. The devices can be both stationary and mobile. Device maintenance may happen in parallel to the actual production process and other communication services performed at the device side without any negative impact on these production communication services.

Table A.2. 5.1-1: Service performance requirements for remote access and maintenance

Use case #	Characteristic parameter				Influence quantity						
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: lower bound	Transfer interval: upper bound	Survival time	UE speed	# of UEs	Service area (note)
1	–	~ 1 month	–	≥ 1 Mbit/s	–	–	–	–	≤ 72 km/h	≤ 100	50 m x 10 m x 10 m
NOTE: Length x width x height.											

Use case one

Transmission of non-deterministic messages in parallel to other interactions. Example applications: software/firmware updates and exchange of maintenance information.

A.3 (void)

A.4 Electric-power distribution

A.4.1 Overview

The energy sector is currently subject to a fundamental change, which is caused by the evolution towards renewable energy, i.e. an increasing number of power plants based on solar and wind power. These changes lead to bi-directional electricity flows and increased dynamics of the power system. New sensors and actuators are being deployed in the power system to efficiently monitor and control the volatile conditions of the grid, requiring real-time information exchange [11][12].

The emerging electric-power distribution grid is also referred to as Smart Grid. The smartness enhances insight into both the grid as a power network and the grid as a system of systems. Enhanced insight improves controllability and predictability, both of which drive improved operation and economic performance and both of which are prerequisites for the sustainable and scalable integration of renewables into the grid and the potential transition to new grid architectures. Smart Grid benefits spread across a broad spectrum but generally include improvements in: power reliability and quality; grid resiliency; power usage optimisation; operational insights; renewable integration; insight into energy usage; and safety and security.

Overviews of (future) electric-power distribution can be found elsewhere in the literature [13][14].

A.4.2 Primary frequency control

Primary frequency control is among the most challenging and demanding control applications in the utility sector. A primary frequency control system is responsible for controlling the energy supply injected and withheld to ensure that the frequency is not deviating more than 0,02 % from the nominal value (e.g., 50 Hz in Europe). Frequency control is based on having sensors for measuring the features in all parts of the network at all points where energy generation or storage units are connected to the grid. At these points, electronic power converters, also known as inverters, are equipped with communication units to send measurement values to other points in the grid such as a frequency control unit, or receive control commands to inject more, or less, energy into the local network.

With the widespread deployment of local generation units, i.e. solar power units, or wind turbines, hundreds of thousands of such units, and their inverters, may have to be connected in a larger power distribution network.

Primary frequency control is carried out in one of three ways:

- 1) Centralised control, all data analysis and corrective actions are determined by a central frequency control unit.
- 2) Decentralised control, the automatic routine frequency control is performed by the individual local inverter based on local frequency values. Statistics and other information is communicated to the central frequency control unit.
- 3) Distributed control, the automatic routine frequency control is performed by the individual local inverter based on local and neighbouring frequency values. Statistics and other information are communicated to the central frequency control unit.

Table A.4.2-1: Service performance requirements for primary frequency control

Use case #	Characteristic parameter			Influence quantity				
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval: target value	Survival time	# of UEs	Service area
1	99,999	TBD	~ 50 ms	~ 100	~ 50 ms	TBD	≤ 100 000	several km ² up to 100 000 km ²

Use case one

Periodic communication service supporting message exchange for primary frequency control.

A.4.3 Distributed voltage control

In the evolution towards 100 % renewable electric power production, the objective of voltage control is to balance the voltage in future low voltage distribution grids connecting local loads and prosumers, as well as energy storage facilities. The aim is to stabilise the voltage as locally as possible, so that decisions and control commands can be issued as quickly as possible. Distributed voltage control is a challenging and demanding control application. Consumer devices rely on having stable voltage levels to operate successfully. When future energy networks rely on thousands of local energy generation units relying mostly on solar and wind power, then it is crucial to stabilise the voltage levels in all segments of the distribution grid. Inverters, or electronic power converters, measure the voltage and power and change the amount of power injected into the grid, and they connect and disconnect end-points from the distribution network.

Distributed control means that the automated voltage control shall be performed by the local voltage control units based on local *and neighbouring* voltage and impedance values. Statistics and other information shall be communicated to the central distribution management system, though.

Table A.4.3-1: Service performance requirements for distributed voltage control

Use case #	Characteristic parameter			Influence quantity				
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Message size [byte]	Transfer interval: target value	Survival time	# of UEs	Service area
1	99,999	TBD	~ 100 ms	~ 100	~ 200 ms	TBD	≤ 100 000	several km ² up to 100 000 km ²

Use case one

Periodic communication service supporting message exchange for distributed voltage control.

A.4.4 Distributed automated switching for isolation and service restoration

A power distribution grid fault is a stressful situation. There are self-healing solutions for automated switching, fault isolation and, service restoration. Furthermore, these solutions are ideally suited to handle outages that affect critical

power consumers, such as industrial plants or data centres. Supply interruptions must be fixed within less than a second for critical power consumers. Automated solutions are able to restore power supply within a few hundred milliseconds.

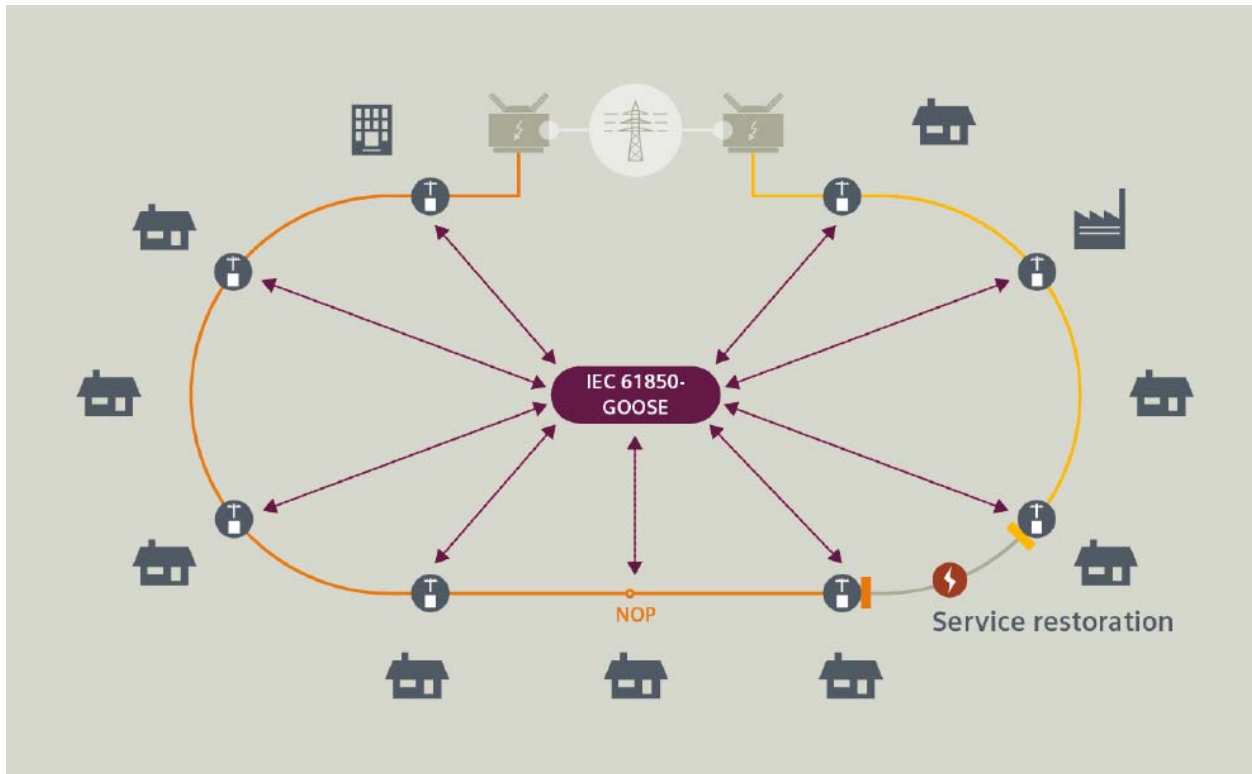


Figure A.4.4-1: Depiction of a distribution ring and a failure (flash of lighting)

The FLISR (Fault Location, Isolation & Service Restoration) solution consists of switch controller devices which are especially designed for feeder automation applications that support the self-healing of power distribution grids with overhead lines. They serve as control units for reclosers and disconnectors in overhead line distribution grids.

The system is designed for using fully distributed, independent automated devices. The logic resides in each individual feeder automation controller located at the poles in the feeder level. Each feeder section has a controller device. Using peer-to-peer communication among the controller devices, the system operates autonomously without the need of a regional controller or control centre. However, all self-healing steps carried out will be reported immediately to the control centre to keep the grid status up-to-date. The controllers conduct self-healing of the distribution line in typically 500 ms by isolating the faults.

Peer-to-peer communication via IEC 61850 GOOSE (Generic Object Oriented Substation Event) messages provides data as fast as possible (Layer 2 multicast message). They are sent periodically (in steady state, with changing interval time in fault case) by each controller to several or all other controllers of the same feeder and are not acknowledged.

The data rate per controller is low in steady state, but GOOSE bursts with high data rate do occur, especially during fault situations. GOOSE messages are sent by several or all controller units of the feeder nearly at the same point in time during the fault location, isolation and service restoration procedure with a low end-to-end latency.

Table A.4.4-1: Service performance requirements for distributed automated switching for isolation and service restoration

Use case #	Characteristic parameter				Influence quantity					
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: target value	Survival time	UE speed	# of UEs	Service area (note 1)
1 (note 2)	99,9999	–	< 5 ms	1 kbit/s (steady state) 1,5 Mbit/s (fault case)	< 1500	< 60 s (steady state) ≥ 1 ms (fault case)	transfer interval (one frame loss)	stationary	20	30 km x 20 km
NOTE 1: Length x width										
NOTE 2: UE to UE communication (two wireless links)										

Use case one

GOOSE (a)periodic deterministic communication service supporting bursty message exchange for fault location, isolation, and service restoration.

A.4.5 Smart grid millisecond-level precise load control

Precise Load Control is the basic application for smart grid. When serious HVDC (high-voltage direct current) transmission fault happens, Millisecond-Level Precise Load Control is used to quickly remove interruptible less-important load, such as electric vehicle charging piles and non-continuous production power supplies in factories.

Table A.4.5-1: Service performance requirements for smart grid millisecond-level precise load control

Use case #	Characteristic parameter				Influence quantity					
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Service bitrate: user experienced data rate	Message size [byte]	Transfer interval: target value	Survival time	UE speed	# of UEs	Service area
1	99,9999	–	< 50 ms	0,59 kbit/s 28 kbit/s	< 100	n/a (note)	–	stationary	10~100/km ²	TBD
NOTE: event-triggered										

Use case one

A non-periodic deterministic communication service between control primary station and load control terminals for removing interruptible less-important load quickly.

A.5 Central power generation

A.5.1 Overview

This domain comprises all aspects of centralised power generation, i.e. the centralised conversion of chemical energy and other forms of energy into electrical energy. Typical electric-power outputs are 100 MW and more. Examples for pertinent systems are large gas turbines, steam turbines, combined-cycle power plants, and wind farms. The planning and installation of respective equipment and plants as well as the operation, monitoring and maintenance of these plants is encompassed by this vertical domain.

A.5.2 Wind power plant network

Table A.5.2-1: Service performance requirements for wind power plant network

Use case #	Characteristic parameter				Influence quantity	
	Communication service availability: target value in %	Communication service reliability: mean time between failures	End-to-end latency: maximum	Packet error ratio	UE speed	Service area
1	99,9999999	~ 10 years	16 ms	< 10 ⁻⁹	stationary	several km ²

Use case one

Communication in support of closed-loop cyber-physical control in a wind farm. The wind farm can be deployed offshore.

NOTE: This type of communication service can be provided via a wired connection.

Annex B (informative): Communication service errors

B.1 Introduction

IEC 61784-3-3 describes fundamental communication errors that can be identified for applications with functional safety requirements [3]. The description of these communication errors is adjusted to field buses. These errors may however also occur in other communication systems. As explained in Annex C, some of these errors are also used for the assessment of communication services that do not support safety-critical applications.

B.2 Corruption

Messages may be corrupted due to errors within an application, due to errors on the transmission medium, or due to message interference.

- NOTE 1: Message error during transfer is a normal event for any standard communication system; such events are detected with high probability at receivers by use of, for instance, hash functions.
- NOTE 2: Most communication systems include protocols for recovery from transmission errors, so these messages will not be classed as 'loss' until recovery or repetition procedures have failed or are not used.
- NOTE 3: If the recovery or repetition procedures take longer than a specified deadline, a message is classed as "unacceptable delay". See also the discussion in Clause C.3.
- NOTE 4: In the very low probability event that multiple errors result in a new message with correct message structure (for example addressing, length, hash function such as CRC, etc.), the message will be accepted and processed further. Evaluations based on a message sequence number or a time stamp can result in fault classifications such as unintended repetition, incorrect sequence, unacceptable delay, insertion [3].

B.3 Unintended repetition

Due to an error, fault, or interference, not updated messages are accidentally repeated.

- NOTE 1: Repetition by the sender is a common procedure when an expected acknowledgment/response is not received from a target station, or when a receiver detects a missing message and asks for it to be resent.

In some cases, the lack of response can be detected, and the message repeated with minimal delay and no loss of sequence, in other cases the repetition occurs later and arrives out of sequence with other messages.

- NOTE 2: Some field buses use redundancy to send the same message multiple times or via multiple alternate routes to increase the probability of good reception [3].

B.4 Incorrect sequence

Due to an error, fault, or interference, the predefined sequence (for example natural numbers, time references) associated with messages from a source is incorrect.

- NOTE 1: Field bus systems can contain elements that store messages (for example FIFOs in switches, bridges, routers) or use protocols that can alter the sequence (for example, by allowing messages with high priority to overtake those with lower priority).
- NOTE 2: When multiple sequences are active, such as transmission of messages from different source entities or reports relating to different object types, these sequences are monitored separately, and errors can be reported for each sequence [3].

B.5 Loss

Due to an error, fault or interference, a message or acknowledgment is not received [3].

B.6 Unacceptable deviation from target end-to-end latency

Messages may be delayed or advanced beyond their permitted arrival time window. Causes for this behaviour include errors in the transmission medium, congested transmission lines, interference, and applications sending messages in such a manner that communication services are delayed or denied.

Message errors can be recovered in the following ways using scheduled or cyclic scans, for instance, in field buses:

- a) immediate repetition;
- b) repetition using spare time at the end of the cycle;
- c) treating the message as lost and waiting for the next cycle to receive the next value.

In case of (a), all subsequent messages in that cycle are slightly delayed, while in case (b) only the resent message is delayed.

Cases (a) and (b) are often not classed as an unacceptable deviation from the target end-to-end latency.

Case (c) would be classed as an unacceptable delay for cyclic, distributed automation functions, unless the cycle repetition interval is short enough to ensure that delays between cycles are not significant and that the next cyclic value can be accepted as a replacement for the missed previous value before the survival time expires (see Clause C.3) [3].

B.7 Masquerade

Due to a fault or interference, a message is inserted that relates to an apparently valid source entity, so a non-safety related message may be received by a safety-related participant, which then treats it as safety related.

NOTE: Communication systems used for safety-related applications can use additional checks to detect masquerade, such as authorised source identities and pass-phrases or cryptography [3].

B.8 Insertion

Due to a fault or interference, a message is received that relates to an unexpected or unknown source entity.

NOTE: These messages are additional to the expected message stream, and because they do not have expected sources, they cannot be classified as correct, unintended repetition, or incorrect sequence [3].

B.9 Addressing

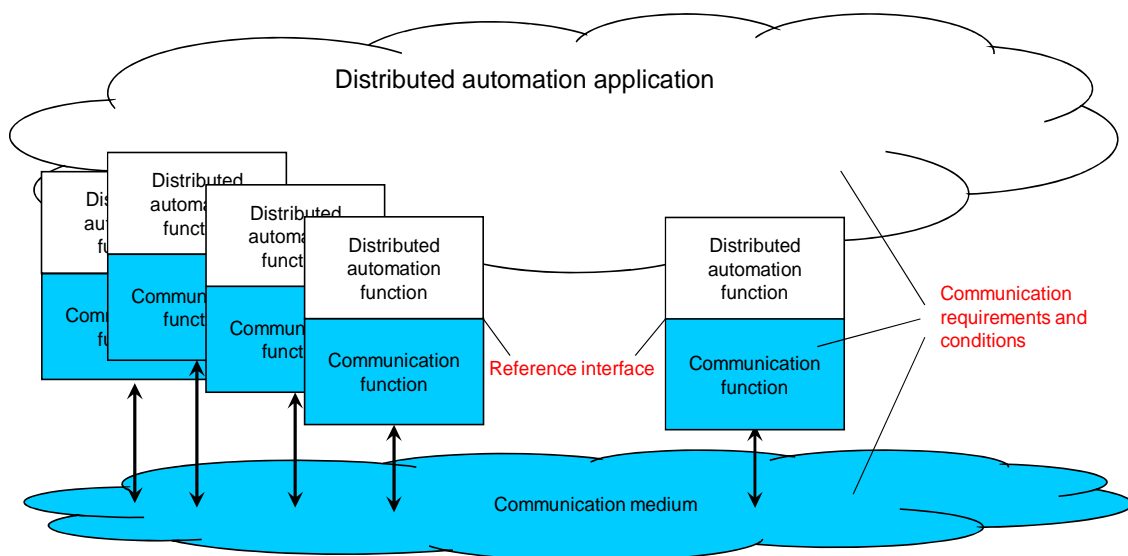
Due to a fault or interference, a safety-related message is delivered to the incorrect safety related participant, which then treats reception as correct [3].

Annex C (informative): Characterising communication services

C.1 Modelling of communication in automation

C.1.1 Area of consideration

For our discussion of communication in automation we apply a definition of the area of consideration for industrial radio communication that is found elsewhere in the literature [4]. This definition is illustrated in Figure C.1.1-1.



NOTE: Blue objects: communication system; other objects: automation application system.

Figure C.1.1-1: Abstract diagram of the area of consideration for industrial radio communication

Here, a distributed automation application system is depicted. This system includes a distributed automation application, which is the aggregation of several automation functions. These can be functions in sensors, measurement devices, drives, switches, I/O devices, encoders etc. all of these functions contribute toward the control of physical objects. Field bus systems, industrial Ethernet systems, or wireless communication systems can be used for connecting the distributed functions. The essential function of these communication systems is the distribution of messages among the distributed automation functions. For cyber-physical control applications, the dependability of the entire communication system and/or of its devices or its links is essential. Communication functions are realised by the respective hardware and software implementation.

In order for the automation application system to operate, messages need to be exchanged between spatially distributed application functions. For that process, messages are exchanged at an interface between the automation application system and the communication system. This interface is termed the reference interface. Required and guaranteed values for characteristic parameters, which describe the behavioural properties of the radio communication system, as well as some influence quantities refer to that interface.

The conditions that influence the behaviour of wireless communication are framed by the communication requirements of the application (e.g., end-to-end latency), the characteristics of the communication system (e.g., output power of a transmitter), and the transmission conditions of the media (e.g., signal fluctuations caused by multipath propagation).

General requirements from the application point of view for the time and failure behaviour of a communication system are mostly related to an end-to-end link. It is assumed in the present document that the behaviour of the link is representative of the communication system as a whole and of the entire scope of the application.

C.1.2 Logical link

C.1.2.1 Nature and function

Starting with the general approach mentioned in Subclause C.1.1, the logical link can be regarded as a possible asset within the area of consideration (see Figure C.1.1-1). The conditions under which its functions are to be performed are vital for the dependability of the automation application system.

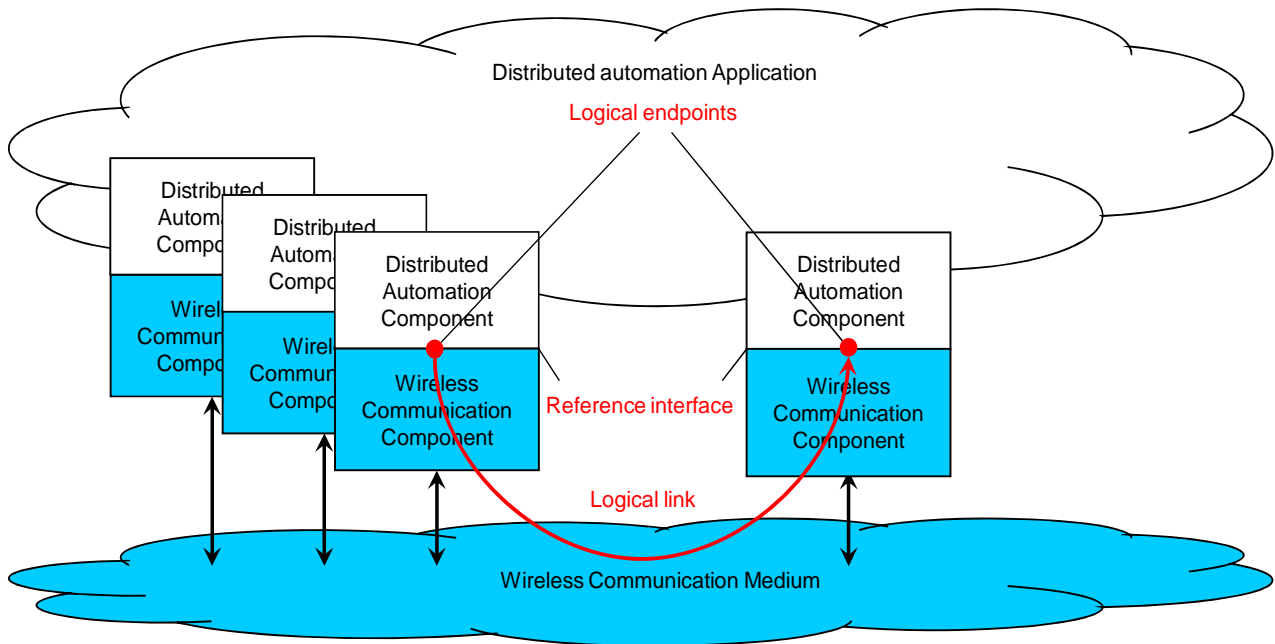


Figure C.1.2.1-1: The concept of a logical link

This is the link between a logical end point in a source device and the logical end point in a target device. Logical end points are elements of the reference interface, which may group several logical end points together.

The intended function of the logical link is the transmission of a sequence of messages from a logical source end point to the correct logical target end point. This is achieved by transforming each message into a form that fosters error-free transmission. The transmission process includes certain processes, for instance repetitions, in order to fulfil the intended function. After transmission, the transported package(s) is converted back into a message. The message is to be available and correct at the target within a defined time. The sequence of messages at the target is to be the same as the sequence at the source.

The functional units, which are necessary to fulfil this function are shown, in Figure C.1.2.1-1.

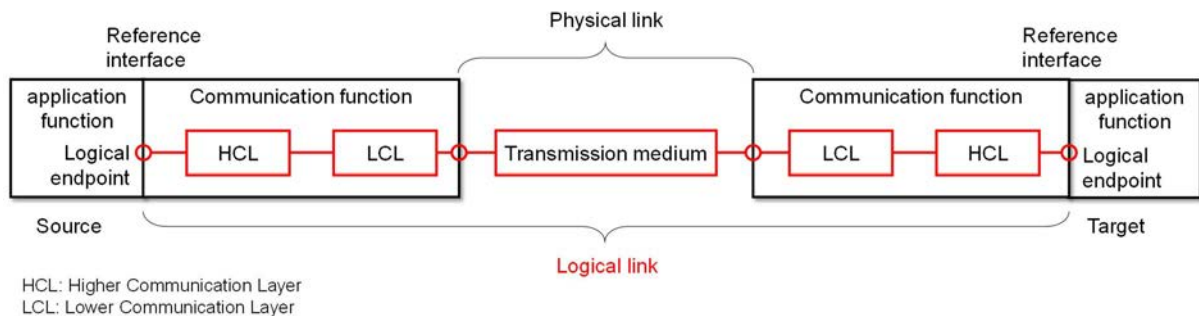


Figure C.1.2.1-2: The asset "logical link"

The required function can be impaired by various influences, which can lead to communication errors. Such errors are described elsewhere in the literature [4][5]. A summary of these errors is provided in Annex A. The occurrence of one of these errors influences the values of the relevant dependability parameters of the logical link.

C.1.2.2 Message transformation

The present document addresses both OSI-layer-3 (IP) and OSI-layer-2 communication. The model in Figure C.1.2-1 can be used for describing both cases. The implementation of communication functions is split between a higher communication layer (HCL) and a lower communication layer (LCL). The partition of the layer for the two traffic options discussed in the present document is provided in Table C.1.2.2-1. This difference is of importance when discussing the implications of the service performance requirements in Clause 5 and Annex A for the network performance (see Clause C.5).

Table C.1.2.2-1: Partition into higher communication layer and lower communication layer

OSI level at which the traffic occurs	Levels comprised by the higher communication layer	Levels comprised by the lower communication layer
3	4 to 6	1 to 3
2	3 to 6 (note)	1 to 2
NOTE: In some vertical application, level 3 to 6 are not implemented.		

The messages to be transmitted for the intended function of a logical link are defined by strings of characters with a certain semantic. Such a character string is handed over as user data at the reference interface for transmission. If the number of characters in a message is too great for it to be transmitted as a unit, the message is divided for transmission into several packets (fragmentation).

C.1.2.3 Communication device

The communication devices—together with the physical link—determine the function and thus the dependability of the logical link. The function of the communication devices is the correct sending and correct receipt of sequences of messages. The asset "communication device" is depicted in Figure C.1.2.3-1.

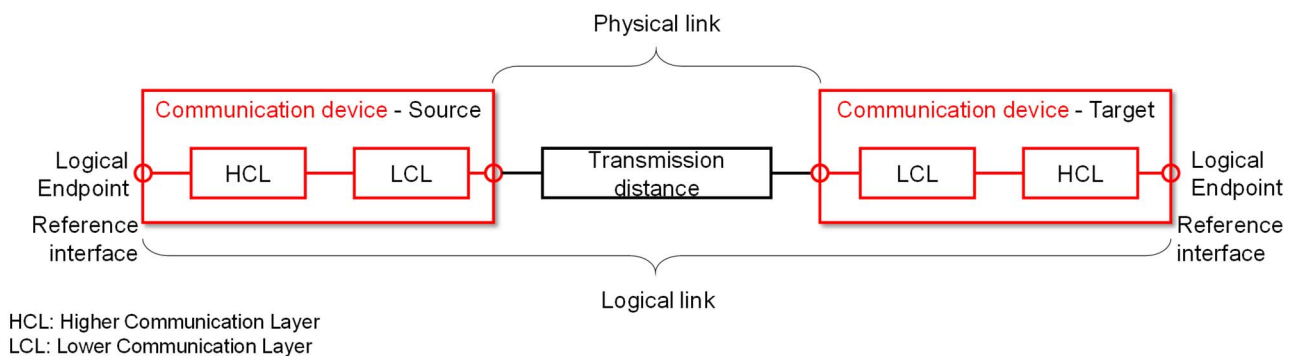


Figure C.1.2.3-1: Asset "communication device"

C.1.2.4 Communication system

The communication system as an asset represents a quantity of logical links whose message transmissions are implemented by wireless devices via one or more media. The communication system function to be provided consists in transmitting messages for all the logical links in the distributed application. This function is to be performed for a defined period, the operating time of the automation application.

In an automation application system, it is paramount that requirements pertaining to logical links are fulfilled. These requirements and the conditions can be very different from one case and implementation to the other. The functions (services and protocols) for individual logical links can therefore also be different. Despite these differences, some of the logical links share communication devices and media.

C.2 Communication service description

C.2.1 Overview

Tables C.2.2-1, C.2.3-1 summarise candidate interface parameters for the description of the communication service performance. The lists are grouped according to whether the parameter stands for automation characteristic parameters (Table C.2.2-1) or influence quantities (Table C.2.3-1). The meaning of the columns and rows is explained after each table.

NOTE 1: Not all parameters in Table C.2.2-1 and Table C.2.3-1 would be used in a service call.

NOTE 2: Ingress and egress in this clause are in reference to the communication service interface between the source application and the communication service interface (ingress) and the communication service and the target application (egress).

C.2.2 Characteristic parameters

Table C.2.2-1: Candidate characteristic parameters for the dependable communication service interface

Parameter name	Typical metric (unit)	Traffic class (note)		
		Deterministic periodic communication	Deterministic aperiodic communication	Non-deterministic communication
Communication service availability	Minimum availability (dimensionless)	X	X	X
End-to-end latency	Target value and timeliness (ms)	X	X	X
Communication service reliability	Mean time between failures (days)	X	X	X
Service bit rate	Target value (bit/s); user experienced data rate (bit/s); time window (s)	–	X	X
Update time	Target value and timeliness	X	–	–

NOTE: – application requirements (KPIs). X: applies; –: does not apply.

Parameter description

Communication service availability

This parameter indicates if the communication system works as contracted ("available"/"unavailable" state). The communication system is in the "available" state as long as the availability criteria for transmitted packets are met. The service is unavailable if the packets received at the target are impaired and/or untimely (e.g. update time > stipulated maximum). If the survival time (see Table C.2.3-1) is larger than zero, consecutive impairments and/or delays are ignored until the respective time has expired.

For the mapping of communication service availability on network KPIs such as packet error rate, it may be assumed that network equipment failures do not happen or only with negligible probability. In this case, only the unavailabilities caused by actual transmission errors (e.g. lost packets, exceeded maximum end-to-end latency) will be considered for the design of network KPIs of the 5G system.

End-to-end latency

This parameter indicates the time allotted to the communication system for transmitting a message and the permitted timeliness.

Communication service reliability

Mean time between failures is one of the typical indicators for communication service reliability. This parameter states the mean value of how long the communication service is available before it becomes unavailable. For instance, a mean

time between failures of one month indicates that a communication service runs error-free for one month on average before an error/errors make the communication service unavailable. Usually, an exponential distribution is assumed. This means, there will be several failures where the time between two subsequent errors is below the mean value (1 month in the example).

Communication service availability and communication service reliability (mean time between failures) give an indication on the time between failures and the length of the failures.

Service bit rate

a) deterministic communication

The target value indicates committed data rate in bit/s sought from the communication service. This is the minimum data rate the communication system guarantees to provide at any time, i.e. in this case target value = user experienced data rate.

b) non-deterministic communication

The target value indicates the target data rate in bit/s. This is the information rate the communication system aims at providing on average during a given (moving) time window (unit: s). The user experienced data rate the lower data rate threshold for any of the time windows.

Update time

Applicable only to periodic communication, the update time indicates the time interval between any two consecutive messages delivered from the egress (of the communication system) to the application.

Traffic classes

In practice, vertical communication networks serve applications exhibiting a wide range of communication requirements. In order to facilitate efficient modelling of the communication network during engineering, and for reducing the complexity of network optimisation, disjoint QoS sets have been identified. These sets are referred to as traffic classes [6]. Typically, only three traffic classes are needed in industrial environments [6], i.e.

- deterministic periodic communication;
- deterministic aperiodic communication; and
- non-deterministic communication.

Deterministic periodic communication stands for periodic communication with stringent requirements on timeliness of the transmission.

Deterministic aperiodic communication stands for communication without a pre-set sending time. Typical activity patterns for which this kind of communication is suitable are event-driven actions.

Non-deterministic communication subsumes all other types of traffic. Periodic non-real time and aperiodic non-real time traffic are subsumed by the non-deterministic traffic class, since periodicity is irrelevant in case the communication is not time-critical.

Usage of the parameters in Table C.2.2-1

Control service request and response; monitoring service response and indication.

C.2.3 Influence quantities

Table C.2.3-1: Candidate application influencing parameters for the dependable communication service interface

Parameter name	Typical metric (unit)	Traffic class (note)			Usage of this parameter
		Deterministic periodic communication	Deterministic aperiodic communication	Non-deterministic communication	
Burst	Maximum user data length (byte) and minimum transfer interval (s)	–	X	–	Control service request and response; monitoring service response and indication
	Message length (byte) and line rate of the service interface (Mbit/s)	–	–	X	
Survival time	Maximum (s)	X	X	–	Control service request and response
Message size	Maximum or current value (byte)	X	(X)	(X)	Control service request and response; non-deterministic data transmission; deterministic aperiodic data transmission
Transfer interval	Target value and timeliness (s)	X	–	–	Control service request and response

NOTE: X: applies; (X): usually does not apply; –: does not apply.

Parameter description

Burst

The transmission of, for instance, program code and configuration data may be handed to the 3GPP system as data burst. In this case, the ingress data rate exceeds the capacity of the network, which implies that some of the data has to be stored within the ingress node of the communication system before it can be transmitted to the egress interface(s). However, the data of a burst needs to be transmitted completely. This is in contrast to periodic data transmission, where new messages overwrite old ones.

Typical metrics for bursts:

- aperiodic data transmission: maximum user data length and minimum transfer interval;
- non-deterministic transmission: message length and line rate of the service interface.

Survival time

The maximum survival time indicates the time period the communication service may not meet the application's requirement before the communication service is deemed to be in an unavailable state.

NOTE 1: The survival time indicates to the communication service the time available to recover from failure. This parameter is thus tightly related to maintainability [7].

Transfer interval

Applicable only to periodic communication, the transfer interval indicates the time elapsed between any two consecutive messages delivered by the automation application to the ingress of the communication system.

Message size

The user data length indicates the (maximum) size of the user data packet delivered from the application to the ingress of the communication system and from the egress of the communication system to the application. For periodic communication this parameter can be used for calculating the requested user-experienced data rate. If this parameter is not provided, the default is the maximum value supported by the PDU type (e.g. Ethernet PDU: maximum frame length is 1522 octets, IP PDU: maximum packet length is 65 535 octets).

C.3 Up time and up state vs. down state and down time

The assessment of periodic deterministic communication services is based on the assessment of successful message transmission over a logical communication link. Message transmission is either:

- *successful, if it is correctly and timely received, or*
- *unsuccessful, if it is incorrectly received, lost or untimely.*

Up time and down time can be derived from received messages. As far as timely received messages are correct, the logical communication link status is *up*. If a message loss or an incorrectly or untimely received message is detected the logical communication link status is *down*. To denote up and down states the terms “up time interval” and “down time interval”, or alternatively “available” and “unavailable” may be used. An example of the relation between logical communication link status, communication service status and application status is presented in Figure C.3-1.

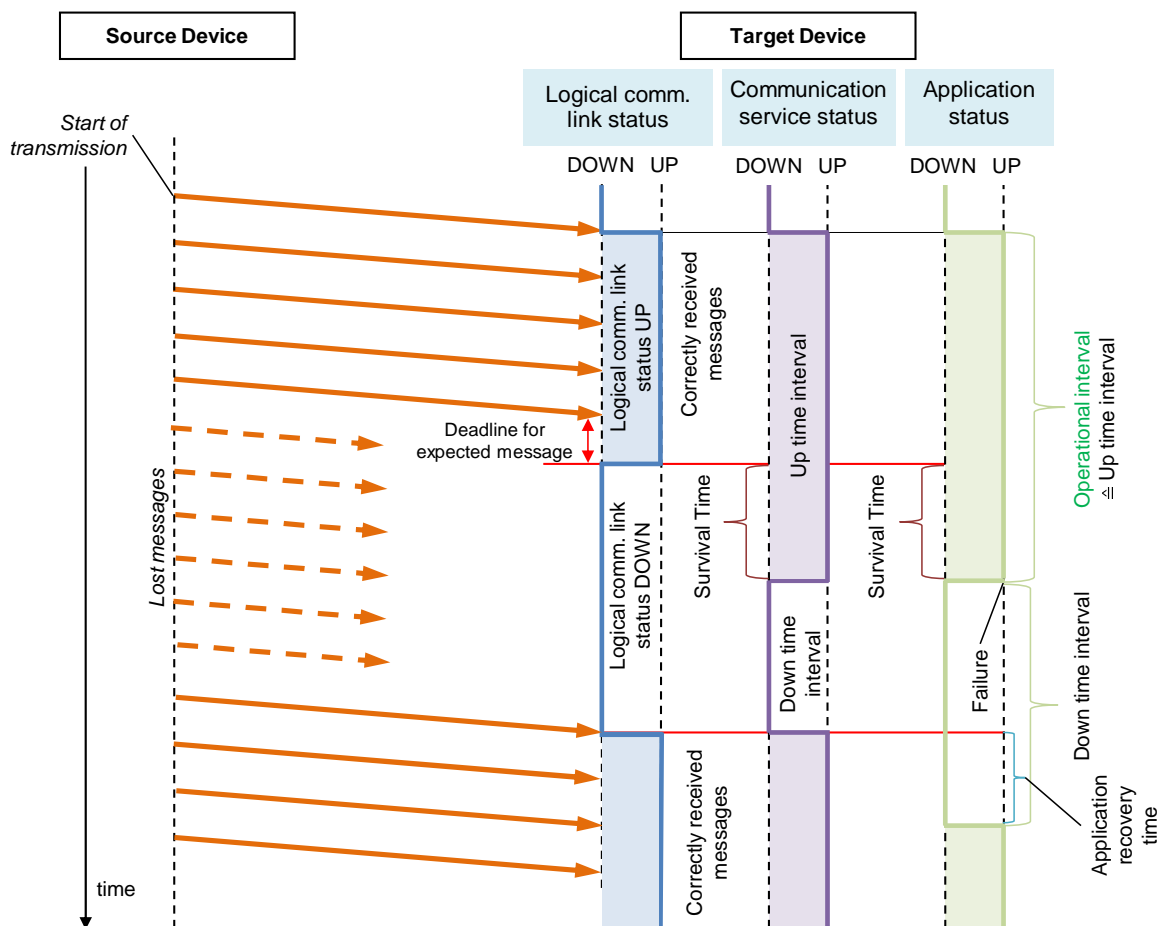


Figure C.3-1: Relation between logical communication link, communication service and application statuses (example with lost messages)

The flow of events in Figure C.3-1 is as follows:

- a) The logical communication link is up and running (blue line is UP). A source device starts sending periodic messages to a target device (orange arrows), on which an automation function (application) is running. The communication service is, from the point of view of the target application, in an up state (violet line is UP) and so is the application (green line is UP).
- b) The logical communication link status changes to down state if it no longer can support end-to-end transmission of the source device's messages to the target device in agreement with the negotiated communication requirements. Once the application on the target device senses the absence (or unsuccessful reception) of expected messages ("Deadline for expected message" in Figure C.3-1), it will wait a pre-set period before it considers the communication service to be unavailable ; this is the so-called survival time. The survival time can be expressed as
 - a period or,
 - especially with cyclic traffic, as maximum number of consecutive incorrectly received or lost messages.
- c) If the survival time has been exceeded, both the communication service and the application transition into a down state (violet and green lines change to DOWN in Figure C.3-1). The application will usually take corresponding actions for handling such situations of unavailable communication services. For instance, it will commence an emergency shutdown. Note that this does not imply that the target application is "shut off"; rather it transitions into a pre-defined state, e.g. a safe state. In the safe state, the target application might still listen to incoming packets or may try to send messages to the source application.
- d) Once the logical communication link status is in the up state again (blue line in Figure C.3-1 changes to UP), the communication service state as perceived by the target application will change to the up state. The communication service is thus again perceived as available (violet line changes to UP in Figure C.3-1). The state of the application, however, depends on the counter measures taken by the application. The application might stay in down state if it is in a safe state due to an emergency shutdown. Or, the application may do a recovery and change to up state again. The time needed for the application to return to the up state after the communication service is restored is shown as "Application recovery time" in Figure C.3-1.

The availability of the communication service is calculated using the accumulated down time. For instance, in case the communication service is expected to run for a time T , the unavailability U of the communication service can be calculated as

$$U = \frac{\sum_i \Delta t_i}{T}$$

Where Δt_i is the length of the i -th downtime interval of the communication service within the time period T . The communication service availability A can then be calculated as

$$A = 1 - U.$$

C.4 Timeliness as an attribute for timing accuracy

C.4.1 Overview

There are several time parameters in dependability assessment. A required value is specified for every time parameter. This value can be a maximum, mean, modal, minimum etc. Typically, there is a deviation from the desired value to the actual value. Jitter is often used to characterise this variation. Since jitter generally is used for characterising the behaviour of a measured parameter, for instance the scatter of measured end-to-end latencies ("the world as it is"), it can be quite confusing to use it for formulating service performance requirements ("the world as we want it to be"). What is needed is a concept and related parameters that allow for formulating and talking about the end-to-end latency requirements in Clause 5 and Annex A.

The most important attribute is timeliness. Timeliness can be formulated a permitted interval for the actual value of the time parameter. Accuracy, earliness and lateness describe the allowed deviation from a target value. Accuracy is the magnitude of deviation. It can be negative (early) or positive (tardy).

C.4.2 Network latency requirement formulated by use of timeliness

In 5G networks, the end-to-end latency KPI is a critical KPI in order to ensure that the network can deliver the packet within a time limit specified by an application: not too early and not too late.

In cyber-physical automation, the arrival time of a specific packet should be strictly inside a prescribed time window. In other words, a strict time boundary applies: [minimum end-to-end latency, maximum end-to-end latency]. Otherwise, the transmission is erroneous. Although most use cases that require timely delivery only specify the maximum end-to-end latency, the minimum latency is also sometimes prescribed. In the latter case, a communication error occurs if the packet is delivered earlier than the minimum end-to-end latency. An example for a related application is putting labels at a specific location on moving objects, and the arrival of a message is interpreted as a trigger for this action. In other words, the application does not keep its own time, but interprets the message arrival as clock signal. Maximum and minimum end-to-end latency alone do not disclose which value is preferred, i.e. target value. The next three subclauses introduce concepts help with relating maximum end-to-end latency, minimum end-to-end latency, and target value to each other.

C.4.3 Timeliness

Timeliness is described by a time interval (see Figure C.4.3-1). The interval is restricted by a lower bound (t_{LB}) and an upper bound (t_{UB}). This interval contains all values t_A that are within an accepted "distance" to the target value t_R .

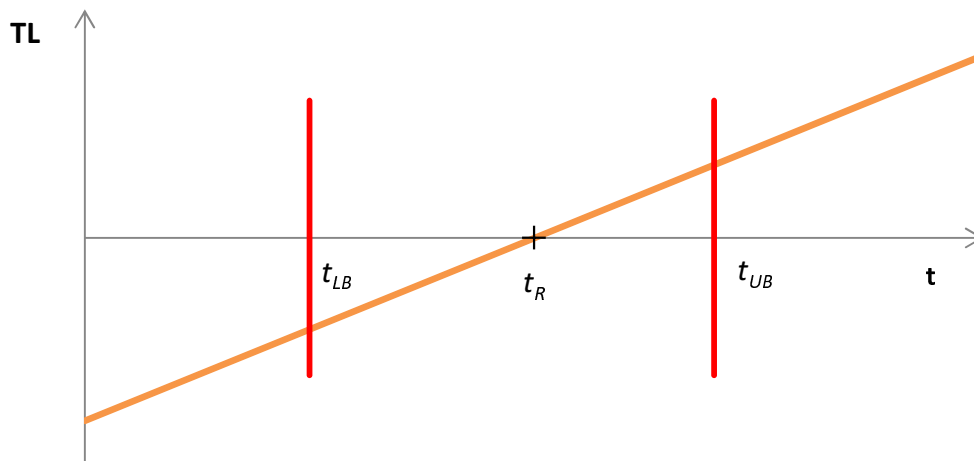


Figure C.4.3-1: Timeliness function

A message reception is considered in time, if it is received within the timeliness interval. If it is received outside the timeliness interval, the message reception is considered invalid. This is related to the communication error "unacceptable deviation from target end-to-end latency" (see Subclause B.6). In other words, maximum end-to-end latency = t_{UB} and minimum end-to-end latency = t_{LB} .

Timeliness is related to deviation (see Subclause C.4.4), the lower bound t_{LB} is related to earliness (see Subclause C.4.5), and the upper bound t_{UB} is related to lateness (see Subclause C.4.6).

C.4.4 Deviation

The term deviation describes the discrepancy between an actual value (t_A) and a target value (t_R).

$$\text{Deviation}(t_A) = t_A - t_R.$$

Figure C.4.4-1 shows two examples. The target value is 10 time units ($t_R = 10$) in both cases. In the first case (blue) the actual value measures 12 time units ($t_A = 12$). The difference of both amounts to +2 time units, which means that the

deviation is 2 time units [$\text{Accuracy}(t_A) = 2$]. The second case (purple) shows the actual value as 9 time units ($t_A = 9$). The difference of both amounts to -1 time unit, which means that the deviation is -1 time units [$\text{Accuracy}(t_A) = -1$].



Figure C.4.4-1: Examples for accuracy values

Figure C.4.4-2 shows the deviation with respect to the target time (t). The following applies:

$\text{Deviation}(t) < 0$ for $t < t_R$; that is, the arrival is early.

$\text{Deviation}(t) = 0$ for $t = t_R$; that is, the arrival is as desired, i.e. on time.

$\text{Deviation}(t) > 0$ for $t > t_R$; that is, the arrival is late (see also C.4.6)

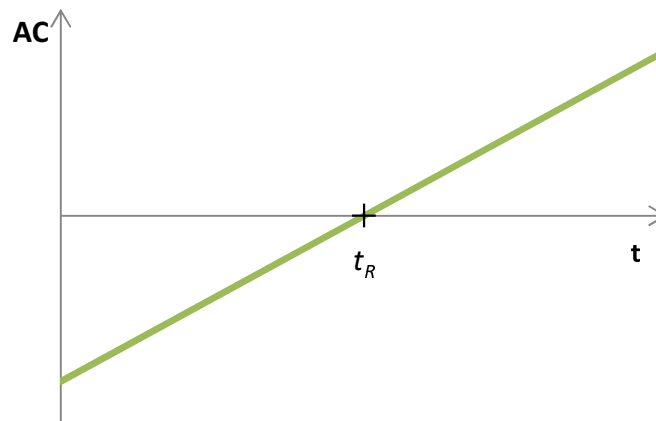


Figure C.4.4-2: Accuracy function

C.4.5 Earliness

Earliness describes how early the actual value is: earliness is greater than 0 if the actual value is less than the target value (see Figure C.4.5-1). The following applies:

$$\text{Earliness}(t_A) = t_R - t_A = -\text{Deviation}(t_A) \text{ for } t_A < t_R;$$

$$\text{Earliness}(t_A) = 0 \text{ for } t_A \geq t_R.$$

In an example, the target value is 10 time units ($t_R = 10$), and the actual value is 7 time units ($t_A = 7$). The difference of both is 3 time units with respect to being early. That means that the earliness is 3 time units [$\text{Earliness}(t_A) = 3$].

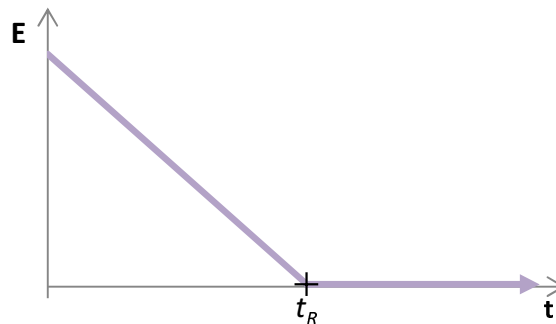


Figure C.4.5-1: Earliness function

C.4.6 Lateness

Lateness describes how much greater the actual value is than the target value: lateness is greater than 0 if the actual value is greater than the desired value (see Figure C.4.6-1). The following applies:

$$L(t_A) = 0 \text{ for } t_A \leq t_R;$$

$$L(t_A) = t_A - t_R = \text{Deviation}(t_A) \text{ for } t_A > t_R.$$

In an example, the target value is 10 time units ($t_R = 10$), and the actual value measures 14 time units ($t_A = 14$). The difference of both is 4 time units with respect to being late. That means that the lateness is 4 time units [$L(t_A) = 4$].

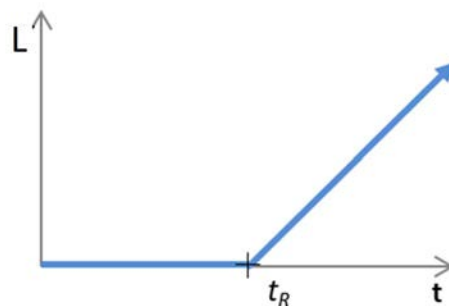


Figure C.4.6-1: Lateness function

C.4.7 Conclusion

Using the concepts of earliness and lateness (see Subclauses C.4.5 and C.4.6, respectively), the maximum and minimum end-to-end latency can be rewritten as follows.

Maximum end-to-end latency = target end-to-end latency + maximum lateness;

Minimum end-to-end latency = target end-to-end latency – maximum earliness.

C.5 Communication service terminology w.r.t. 5G network and vertical applications

This section clarifies the wording and terminology with respect to communication interfaces that are relevant for vertical applications. Because the 3GPP network does not cover the complete ISO-OSI communication stack, it is important to distinguish between

- the vertical applications' point of view, and
- the 3GPP network's point of view.

In this section, the relation between those two is clarified.

Figure C.5-1 shows a simplified version of the communication stack. The PHY layer, the MAC layer and some parts of the IP layer are part of the 3GPP network. The layers that are part of the 3GPP network are referred to as lower communication layers (LCL). The communication stack also includes an application. The OSI layers related to providing data to the application are referred to as the higher communication layers (HCL). The interface between LCL and HCL is referred to as communication service interface (CSIF).

For the assessment of the overall system performance, it is important to differentiate between the 3GPP network's performance (i.e., including only the LCL and measured at the CSIF) and the overall system performance including the application layer (i.e., including both, the LCL and the HCL). In Figure C.5-1, the orange arrow depicts the vertical application's point of view. The blue arrows indicate two options to measure the 3GPP network's performance, i.e., including and excluding the IP layer.

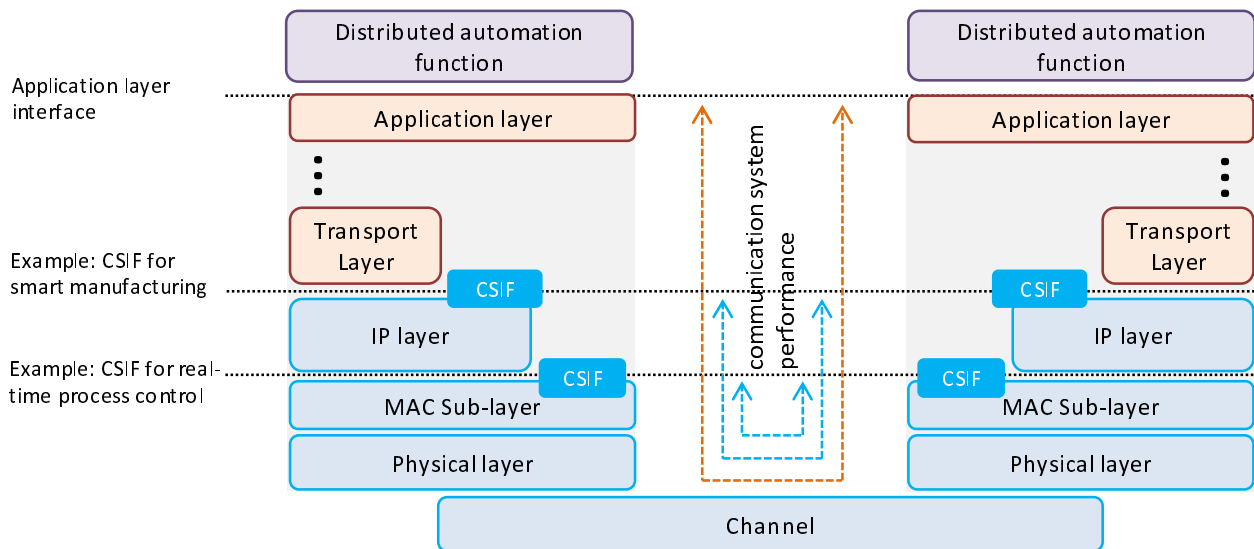


Figure C.5-1: Network performance measurements at different communication system interfaces (CSIF)

Figure C.5-2 illustrates how messages are transmitted from a source application device (e.g., a programmable logic controller) to a target application device (e.g. an industrial robot). The source application function (AF) is executed in the source operating system (OS) and hands over a message to the application layer interface of the source communication device. In the higher communication layers (HCL), which are not part of the 3GPP system, the data is processed. From the HCL the data is transferred to the lower communication layers (LCL), which are part of the 3GPP system. After transmission through the physical communication channel and the LCL of the target communication device, the data is passed to the HCL and lastly to the target application device. Characteristic parameters with respect to time are defined in Figure C.5-2.

From 3GPP system point of view:

- Transfer interval of 5G system: Time between the arrival of two pieces of data at the source CSIF.
- End-to-end latency: Time measured from the point when a piece of data received at the CSIF in the source communication device until the same piece of data is passed to the CSIF in the target communication device.

From vertical application point of view:

- Transfer interval of vertical application: Time between the transmission of two successive pieces of data from the source application.
- Transmission time: Time measured from the point when a piece of data is handed from the application layer interface of the source application device, until the same piece of data is received at the application layer interface of the target application device.
- Update time: Time between the reception of two consecutive pieces of data at the application layer interface to the target application device.

If not stated otherwise, the terms "end-to-end latency" and "transfer interval" refer to the 3GPP system / 5G network parameters in this document.

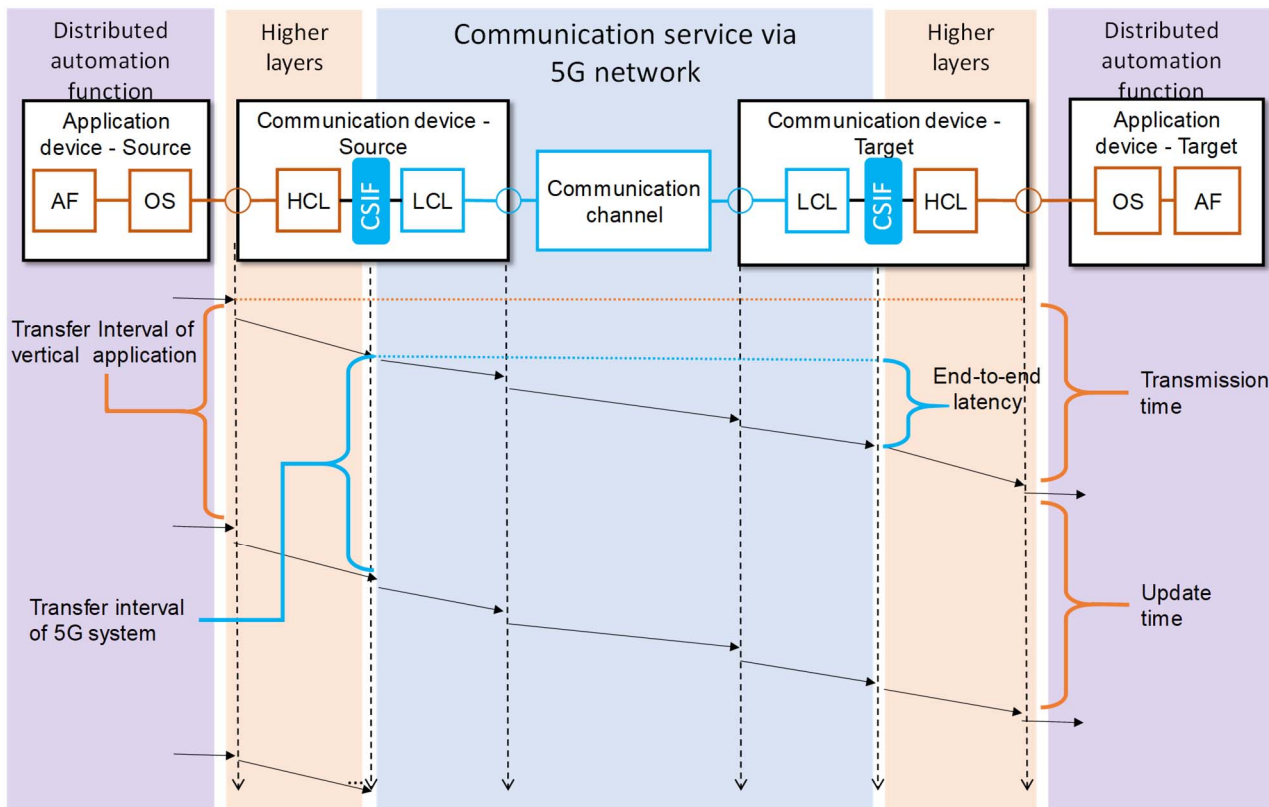


Figure C.5-2: Relation between application device and communication device (downlink example).

Annex D (informative): 5G in industrial automation: different and multiple time domains for synchronization

D.1 Description

The required synchronization precision is usually given as the maximum absolute value of the time difference between sync master and any device in the synchronisation domain (time domain or clock domain). A common example is a synchronisation precision of $\leq 1 \mu\text{s}$. This is equivalent to \pm the precision value, so $\pm 1 \mu\text{s}$ between sync master and any device in the synchronisation domain, resulting in two times this value as maximum absolute time difference between any two devices in the synchronisation domain ($2 \mu\text{s}$ in the example).

An industrial automation network generally consists of two distinct time domains.

First is the *global time domain*. This is the time used for overall synchronization in the system (e.g. the factory). It is used to align operations and events chronologically. Industrial automation uses the term *universal time domain* [20] for the global time domain described in this document. Global time is known as a synonym for universal time in industrial automation. Global time is called wall clock in certain areas and standards.

The synchronization precision is typically $\leq 1 \mu\text{s}$ [20]. In some areas, a precision of $\leq 100 \mu\text{s}$ might be sufficient for the global time domain if a working clock with precision of $\leq 1 \mu\text{s}$ is available. The assigned timescale is usually the International Atomic Time (TAI, *temps atomique international*), based on the precision time protocol (PTP) epoch (starting from 1 January 1970 00:00:00 TAI) [22]. While there is usually only one global time, multiple global time domains are possible.

Clock synchronization in the global time domain usually applies to all UEs within the industrial facility in industrial automation. That is, a global time domain covers usually the industrial facility.

Second is the *working clock domain*. Working clock domains are constrained in size. They often consist of a single machine or a set of neighbouring machines that physically collaborate. The restricted size allows very precise time synchronization ($\leq 1 \mu\text{s}$) with efficient network components. Synchronisation to a working clock is used to align e.g. production lines, production cells, or machines/functional units. In these cases, the application synchronizes locally within the working clock domains (Figure D.1-1), allowing precise synchronization with more efficient components. A global time domain usually contains multiple working clock domains. The starting point (epoch) is the start of the working clock domain.

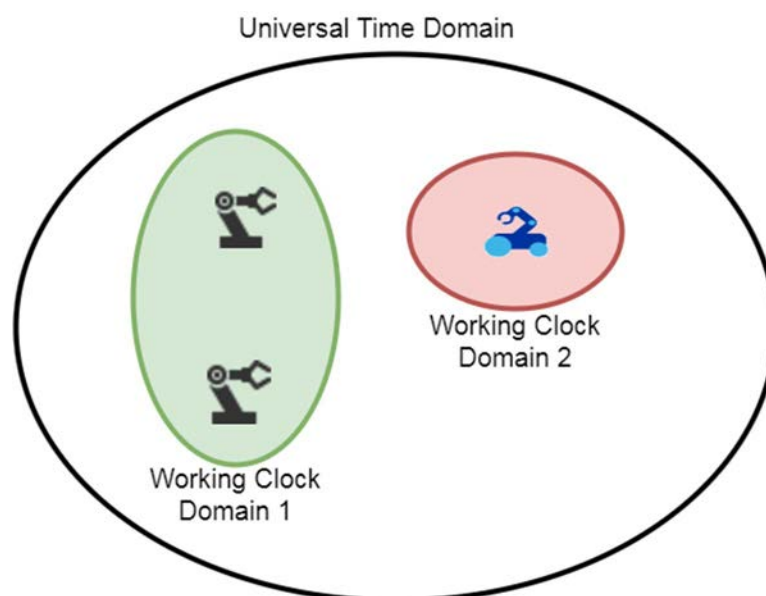


Figure D.1-1: Global time domain and working clock domains

The assigned timescale of a working clock domain is arbitrary (timescale ARB [22]). Therefore, different working clock domains may have different timescales and different synchronisation accuracy and precision. Robots, motion control applications, numeric control, and any kind of clocked / isochronous application rely on the timescale of the working clock domain to make sure that actions are precisely interwoven as needed.

Clock synchronization in the working clock domain is constraint in size. A specific working clock domain will contain only a subset of the UEs within the industrial facility. Often, the UEs of the working clock domain are connected to the same gNB. However, it is also possible that a working clock domain contains multiple neighbouring gNBs. This depends on the actual use case and its vertical application.

Devices may be part of multiple time domains leading to overlapping working clock domains.

The required precision (usually $\leq 1 \mu\text{s}$) is between the sync master and any sync slaves (device) of the clock domain, both, global time domain and working clock domain.

Clock domains might be called sync domains in certain areas and standards.

D.2 Merging of working clock domains

One key issue of the integration of TSN and 5G wireless networks that has to be handled is mobility. The integration of 5G wireless communication into the industrial communication infrastructure allows for mobility in the manufacturing process. This mobility enhances flexibility in the manufacturing process, e.g. through adding certain manufacturing capabilities on-demand by having a machine move to the respective production line. This means that machines that are synchronized to different working clock domains may need to interact with each other.

The following scenario illustrates this. After the mobile machine has arrived at the intended location and is stationary again, the two interacting working clock domains have to be synchronized with each other. Otherwise interaction might not be possible without interfering with ongoing operations. An example is an autonomous mobile handling robot adding parts to an assembly line. Without synchronization between both, correct placement of the parts would be impossible.

However, it is not feasible to schedule these interactions beforehand. Therefore, the interaction between different working clock domains requires a concept for handling the communication. TSN provides already mechanisms for this. The 5G systems and the UEs need to provide an interface in order to exchange information of the clock domain.

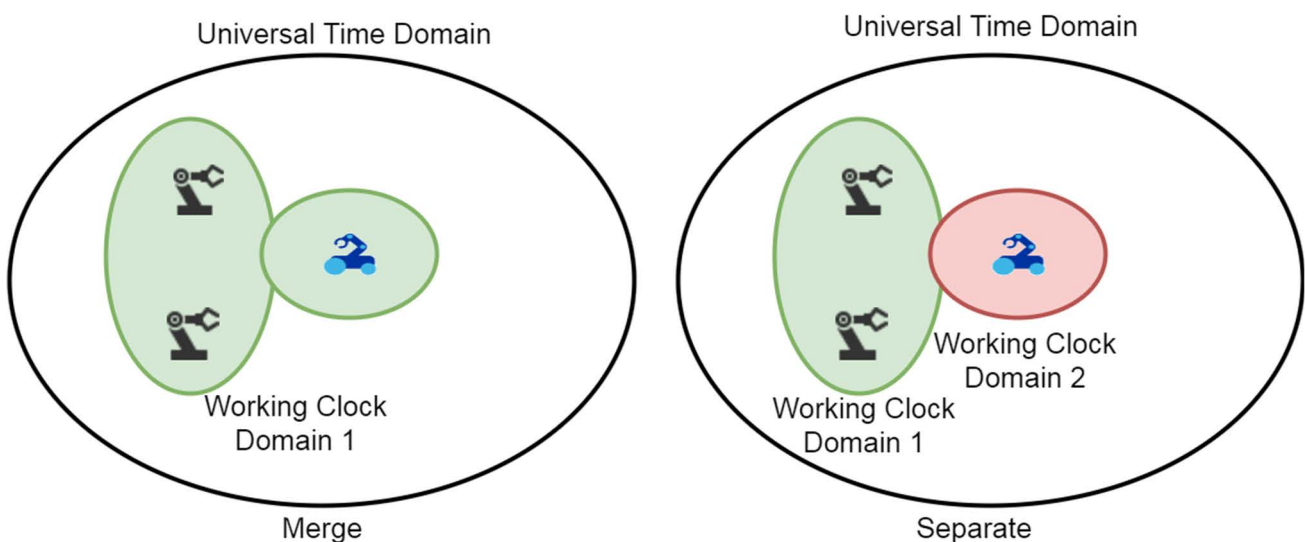


Figure D.2-1: Working clock domain interactions "Merge" and "Separate"

When members of different working clock domains interact, there are two possible options (Figure D.2-1). Which option is used depends on the application and its requirements.

- Merge: The working clock domains merge into one. This option can be used in applications where synchronization is critical, e.g. high precision robots interacting with each other.

- Separate: The members of the different working clock domains interact while keeping their own separate time synchronizations. This option can be used in applications where synchronization is non-critical, e.g. an AGV collecting finished products from a production line.

D.3 Time synchronization with 5G networks

For the time synchronization with 5G networks, we consider two possible options.

The 5G system uses the 802.1AS time sync domains: In this case, the 5G system provides a media dependent interface to the 802.1AS sync domain [21], which the application can use to synchronize to the sync domain. In the IEEE 802.1AS standard [22], a similar concept is detailed in the MDSyncSend and MDSyncReceive structures.

The 5G system provides the working clock domains and global time domain: In this case, the 5G system has to provide an interface which the application can use to derive their working clock domain or global time domain. A device can belong to multiple working clock domains. An application can use each of these as the reference clock for synchronization (reference clock model).

NOTE: The required precision (usually $\leq 1 \mu\text{s}$) is between the sync master and any sync slave (device) of the clock domain.

Annex E (informative): Change history

Change history							
Date	Meeting	TDoc	CR	Rev	Cat	Subject/Comment	New version
2018-05	SA1#82	S1-181551	–	–	–	Skeleton for TS 22.104 ("Service requirements for cyber-physical control applications in vertical domains")	0.0.0
2018-05	SA1#82	S1-181552	–	–	–	Includes agreements at SA1#82, Dubrovnik, Croatia	0.1.0
2018-08	SA1#83	S1-182344	–	–	–	Includes agreements at SA1#83, West Palm Beach, Florida	0.2.0
2018-11	SA1#84	S1-183276	–	–	–	Includes agreements at SA1#83, Spokane, WA, USA, rapporteur's clean-up	0.3.0
2018-12	SA#82	SP-181006				Presentation to SA for one-step approval	1.0.0
2018-12	SA#82	SP-181006				Raised to v.16.0.0 following SA approval	16.0.0
2019-03	SA#83	SP-190081	0003	2	F	Clarifying UE-to-UE versus UE-to-network	16.1.0
2019-03	SA#83	SP-190081	0002	1	F	Moving rail-bound mass transit requirements – shift from cyberCAV	16.1.0
2019-03	SA#83	SP-190081	0001	1	F	Clean-up and corrections of TS 22.104 cyberCAV	16.1.0
2019-06	SA#84	SP-190299	0008		F	Corrections to TS 22.104 v16.1.0	16.2.0
2019-06	SA#84	SP-190299	0006	2	F	Add missing abbreviations to TS 22.104	16.2.0
2019-06	SA#84	SP-190299	0005	2	C	Adding edge computing aspect	16.2.0
2019-09	SA#85	SP-190800	0022	1	D	Correction of a figure number in Annex D.2	16.3.0
2019-09	SA#85	SP-190800	0014	3	F	Clarification for CSA requirements	16.3.0
2019-12	SA#86	SP-191028	0033	5	F	Clarification of clock synchronicity requirements	16.4.0
2019-12	SA#86	SP-191028	0029	1	F	Clarification on communication service reliability	16.4.0
2020-07	SA#88e	SP-200562	0051		F	Miscellaneous values for further study	16.5.0
2020-07	SA#88e	SP-200562	0040	2	F	Clarifications to communication service performance requirements	16.5.0
2020-07	SA#88e	SP-200562	0044	2	F	Correcting description of communication service status in Clause C.3	16.5.0
2020-07	SA#88e	SP-200562	0046	1	F	Clock synchronicity budget for the 5G system	16.5.0

History

Document history		
V16.5.0	September 2020	Publication