

Report

5G Common Terminology

Grant agreement number:	857008
Project title:	5G Smart Manufacturing
Project acronym:	5G-SMART
Project website:	www.5gsmart.eu
Programme:	H2020-ICT-2018-3
Contributing workpackages:	WP2, WP3, WP4, WP5, WP6
Dissemination level:	internal
Responsible organization:	Orange
Editor(s):	Berna Sayrac
Version number:	0.5
Status:	Draft
Short abstract/summary:	The aim of this document is to clarify the key concepts, definitions and terms used by OT and ICT partners throughout the project so that we have a common understanding on the language used during the project discussions and deliverables. This is a living document that is expected to evolve during the project lifetime as new needs of clarification and common understanding arises.
Keywords:	5G, Factory of the Future, Industry 4.0, IoT, Factory Automation, Process Automation, Industrial Robotics, Cloud Robotics

Contributor(s):	(ordered according to beneficiary number) Leeke Grosjean (ERI-SE) Milad Ganjalizadeh (ERI-SE) Juan-Antonio Ibanez (ERI-SE) Ognjen Dobrijevic (ABB) Krister Landernäs (ABB) Ahmad Rostami (BOSCH) Niels König (IPT) Gabor Nemeth (ERI-HU) Roberto Padovani (MARPA) Berna Sayrac (ORANGE) Fanny Parzysz (ORANGE) Giyyarpuram Madhusudan (ORANGE) Guillaume Boulmier (ORANGE) Olivier Le Moult (ORANGE) Saul Inca (UPV)
-----------------	---



Fredrik Tufvesson (ULUND)

Maliosz Markosz (BME)

Kimmo Hiltunen (ERI-FI)

Contents

1	Introduction	3
1.1	Reference framework	3
1.2	Edge Computing.....	4
1.3	Metrics and Key Performance Indicators (KPIs).....	8
1.3.1	Data rate	8
1.3.2	Timeliness aspects	9
1.3.3	Dependability.....	9
1.4	Communication Characterization	10
1.4.1	Geographical characteristics.....	10
1.4.2	Traffic models.....	11
1.5	Production/Manufacturing related terms	12
2	References	13
3	Appendix	15
3.1	Derivation of communication service availability and reliability from network performance metrics	15
3.1.1	Additional definitions.....	15
3.1.2	Mapping function between communication service and network.....	16
3.1.3	Network vs Communication service performance	17
	List of abbreviations	19

1 Introduction

5G-SMART brings together a consortium consisting of (Ericsson, Cumucore, T-systems Hungary), network operators (Orange), providers of wireless communication technologies and components (u-blox), operational technologies' suppliers (ABB, Bosch, Fraunhofer IPT, Marposs), factory operators (Bosch) and academia (Lund University, University of Valencia, Budapest University of Technology and Economics). Given the diversity in the background of the consortium members, finding a common language and understanding of definitions and terms has not always been easy and often shown to be a challenge. The work devoted to this important task has led to the development of this document, which is hopefully useful to other collaborations working in the area of smart manufacturing with a similar diversity in the consortium.

The aim of this document is to clarify the key concepts, definitions and terms used throughout the project so that all partners have a common understanding on the language used during the project discussions and deliverables independent of their background. This is a living document that is expected to evolve during the project lifetime as new needs of clarification and common understanding arises.

Note that for most of the terms, we have adopted the definitions that are widely used and accepted in the ecosystem, and for some others we have proposed modifications to the existing ones or proposed totally new ones. Some existing definitions come from the NGMN 5G White Paper [NGMN15] as well as the 3GPP documents TS 22.261 [TS22.261] and 3GPP TS 22.104 [TS22.104]. The reference framework (Figure 1) is mostly based on the NGMN document “Verticals URLLC Use Cases and Requirements” [NGMN2019] but adopted for a 5G smart manufacturing context. For data traffic models, the 5G-ACIA white paper on traffic models [5GACIA] is used, and for Mobile Edge Computing, ETSI documents are taken as a basis.

1.1 Reference framework

An **application service** represents an application that is provided within a use case and scenario with particular characteristics in terms of key performance indicators (KPIs). Examples of an application service are Augmented Reality (AR), Virtual Reality (VR), Ultra-High Definition (UHD) video, real-time control, etc.

To realize an application service that involves distinct connected objects, a **communication service** must be set up and may need to fulfill specific requirements to allow proper operations of these connected objects. Principal requirements and performance metrics are explained in Subsection 2.3.3 and the different traffic flow types of such communication services are described in Subsection 2.3.4.

A communication **service category** represents a set of communication services that share some common characteristics in terms of connectivity. eMBB, mMTC and URLLC are examples of service categories. The services within each category may not have the same KPIs but share some basic connectivity parameters, like, e.g., the need to provide reliability and low latency in URLLC.



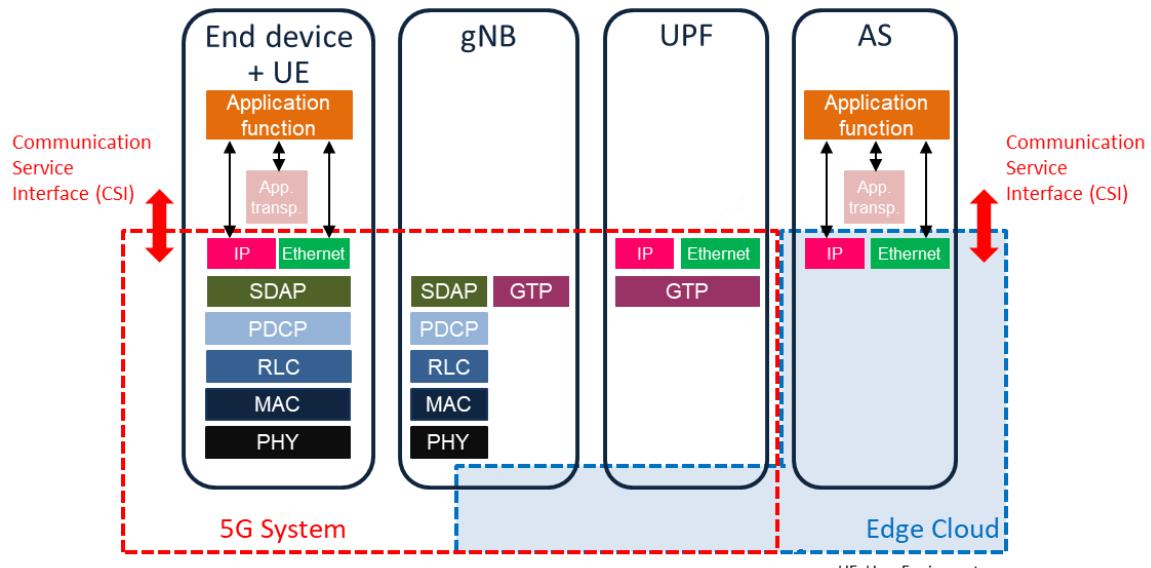
Two fundamental perspectives are considered to evaluate the performance of a practical 5G system deployment [TS22.104]: i) the **application perspective** of the communication services, which is related to the application using the communication service (e.g., a robot motion controller) and ii) the **network perspective**, which sustains this communication service. In between, the **Communication Service Interface (CSI)** connects the application system to the underlying 5G system, as depicted in Figure 1, and serves as the principal reference point for defining metrics and evaluating performance. In this figure,

- **Application Server (AS)** refers to a server specifically designed to run applications. The "server" includes both the hardware and software that provide an environment for programs to run.
- **User Equipment (UE)** refers to the 5G functionality in a (mobile) end device that provides the connectivity to the 5G network.
- **Application function** refers to the functionality that realizes/enables the application service.
- **Application transport** refers to different protocol options for transferring data units of an application service in an end-to-end manner, such as TCP/UDP over IP, and TCP/UDP over IP and over Ethernet.

For most of the smart manufacturing deployments, the UE is not embedded in the end device (like in a smart phone) but connected via a separate L3/L2 connection (there may even be several end devices behind a 5G UE). For those cases, the reference framework looks like in Figure 2.

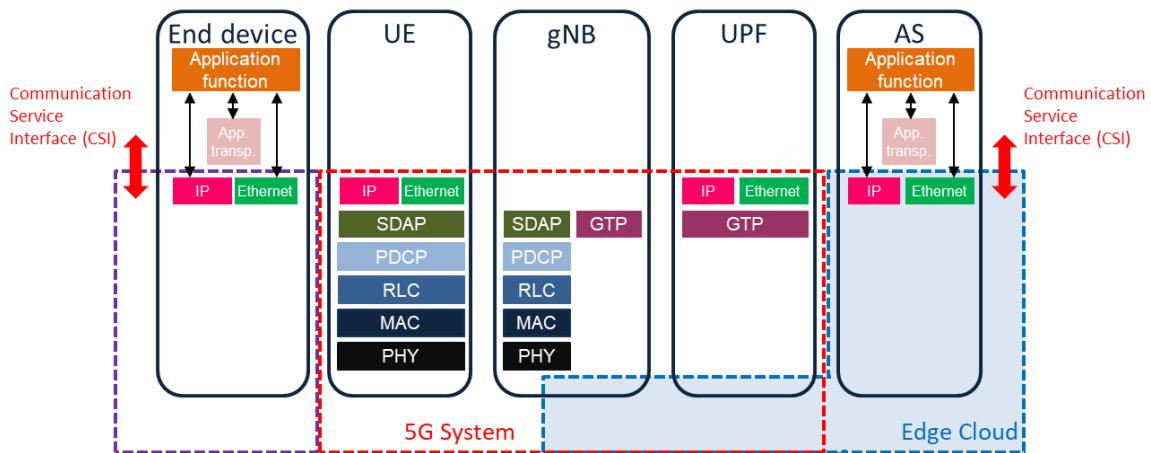
1.2 Edge Computing

Edge computing is a generic term encompassing a variety of different approaches to putting computing and storage resources at the edge of the network close to the customer rather than in remote data centers. It includes approaches like fog computing, cloudlets and others, which are not based on the cellular network. As illustrated in Figure 2, each segment of Edge computing has its own characteristics, potentially addresses different requirements and comes at a different cost, as distributing resources may significantly increase the complexity of integrating, operating and maintaining the overall infrastructure.



UE: User Equipment
gNB: gNodeB (5G radio base station)
UPF: User Plane Function
AS: Application Server
PHY: Physical Layer
MAC: Medium Access Control
RLC: Radio Link Control
PDCP: Packet Data Convergence Protocol
SDAP: Service Data Adaptation Protocol
GTP: GPRS (General Packet Radio Service) Tunneling Protocol

(a)



UE: User Equipment
gNB: gNodeB (5G radio base station)
UPF: User Plane Function
AS: Application Server
PHY: Physical Layer
MAC: Medium Access Control
RLC: Radio Link Control
PDCP: Packet Data Convergence Protocol
SDAP: Service Data Adaptation Protocol
GTP: GPRS (General Packet Radio Service) Tunneling Protocol

(b)

Figure 1. Reference framework of communication systems in smart manufacturing: (a) embedded UE, (b) UE over a L3/L2 connection

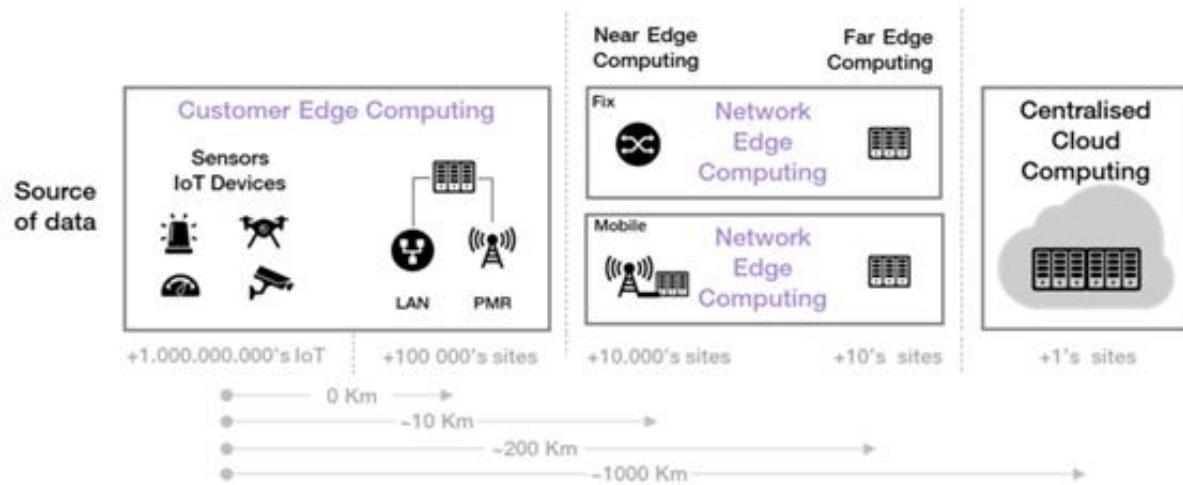


Figure 2. Different types of edge computing (source: Orange)

Initially, this notion was introduced and used for mobile networks, hence the term **Mobile Edge Computing**. Later, the European Telecommunications Standards Institute (ETSI) defined the term **Multi-access Edge Computing (MEC)** as a generalization of Mobile Edge Computing to any network. Although the concept of Edge computing applies to any type of network, in the context of 5G-SMART, the term **Edge Computing** is used to refer to a mobile network architecture model that enables a business oriented, cloud computing platform within the radio access network at the close proximity of the users/devices of the mobile network to serve delay-sensitive, context-aware applications. It also enables Network Function Virtualization. Those edge-based capabilities can be provided to internal network functions, in-house applications run by the operator or the network customers, or potentially third-party partners / developers (Figure 3). ETSI White Paper No. 28 on “MEC in 5G networks” [ETSI5GMEC] offers a good overview of the model.

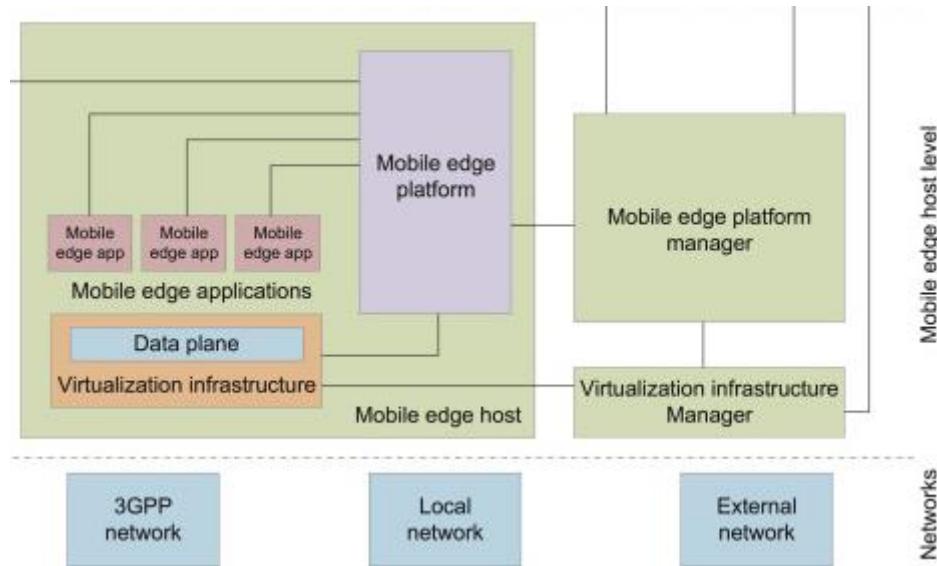


Figure 3. Integrated MEC deployment in 5G network [ETSI5GMEC]

Different Edge compute deployments are possible (e.g. factory cloud, on-premise Edge cloud, telco Edge cloud etc.), the details of which are out of the scope of this document. For more details on the deployment options, one can refer to [ETSI5GMEC] or [SKTW20]. For illustration purposes, a mobile Edge host could be a virtual machine on the same hardware as a 5G radio base station (gNodeB) and a mobile Edge app could be a robot controller software running on this host (Figure 3).

Another model for managing distributed cloud environments in mobile networks, from central cloud computing to edge cloud computing, is defined by the **Open Networking Automation Platform** [ONAP], which is an open source project with wide industry support [ONAP-WP]. ONAP allows to orchestrate and manage network functions and application functions in a distributed computing environment. This enables to provide flexible realizations for different services where applications are embedded into the communication infrastructure according to network characteristics and service needs [BSK18] [SKTW20], see Figure 4.

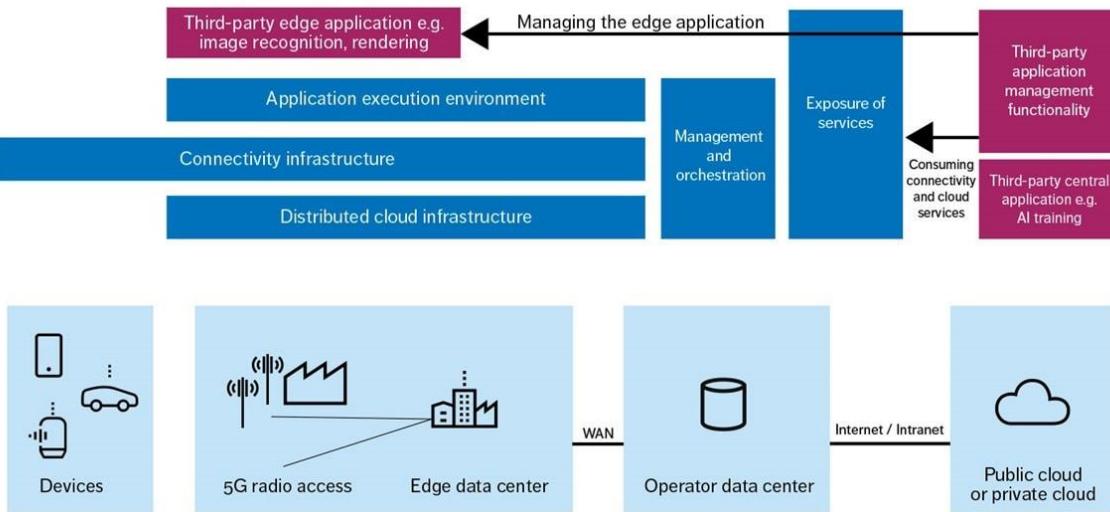


Figure 4. High-level architecture of an edge-computing solution (source: [SKTW20])

NOTE: **Edge cloud** is another term for edge computing, not necessarily only on MEC as defined above but also on other edge architectures. In 5G-SMART, the term **Edge cloud** is used synonymously with **Edge computing**.

Local BreakOut (LBO): defines local termination of the 3GPP-standardized user-plane connectivity close to the base station. This enables to locate time-critical applications in proximity, for example in an edge cloud, and provide access to advanced 5G capabilities (e.g., 5G support for LAN and TSN connectivity) at the local site. LBO combined with 5G URLLC RAN support enables bounded low latency for control loops on the factory shop floor that connect machines and sensors.

1.3 Metrics and Key Performance Indicators (KPIs)

Data rate, timeliness and dependability are key parameters to evaluate the performance of a communication service for smart manufacturing. Related KPIs are further explained in the following.

1.3.1 Data rate

The **user data rate** is defined as the minimum value of the number of bits transmitted or received over time, typically expressed in Mbit/s, which is expected to be measured at the CSI shown in Figure 1. This definition excludes scenarios for broadcast-like services, where the given value is the maximum that is needed [TS 22.261].

For uneven or bursty traffic, the **average (respectively peak) data rate** is defined as the average (respectively maximum) number of bits transmitted or received at the CSI over a time window. The



specific size of this time window depends on the communication service characteristics (such as traffic, performance requirements) and the application that uses this communication service.

1.3.2 Timeliness aspects

Timeliness aspects are essential when it comes to URLLC communication or to synchronized industrial processes. Latency, jitter and synchronization accuracy are defined as follows:

- **(End-to-end) Latency** is the time that it takes to transfer application data of a given size from a source to a destination, from the moment it is transmitted by the source to the moment it is successfully received at the destination (one-way latency). In other words, the end-to-end (E2E) latency, typically expressed in ms, is measured from the CSI on the UE side to the same interface on the application server side (Figure 1), or vice versa.
- **Jitter** is the variation of a time parameter, typically the end-to-end latency. The jitter value is specified by the application service.
- **Time synchronization error** is defined as the value of the time difference between a sync master (that is used as the timing reference) and any device operating on time-sensitive applications. For example, a time synchronization error of $\leq 1 \mu\text{s}$ is equivalent to having a time difference equal to at most $\pm 1 \mu\text{s}$ offset between the sync master (for example global positioning system, GPS) and any device in an industrial network, resulting in two times this value as the maximum absolute time difference between any two devices in the network (i.e., $2 \mu\text{s}$).

1.3.3 Dependability

Dependability involves the following components:

- The **survival time** indicates the maximum time period the communication service may not meet the application's requirements before there is a failure on the application layer, such that the communication service is deemed to be in an unavailable state. Such a situation occurs when the communication with the network is lost, the application crashes and an alarm is raised.
- In the context of network layer packet transmissions, **reliability** is the percentage value of the amount of sent network layer packets successfully delivered to a given system entity within the time constraint required by the targeted service, divided by the total number of sent network layer packets [TS 22.261]. In order to well differentiate reliability from communication service reliability (defined below), we will use the term **network packet transmission reliability**.
- The **communication service availability** relates to the ability to allow correct operation of the application. It is defined as the “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” [TS 22.104]. The service is unavailable if the messages received at the target are impaired and/or untimely (e.g. latency $>$ stipulated maximum), resulting in survival time being exceeded.
- The **communication service reliability** relates to the ability to continuously operate as required by the application, without failure, for a given time interval and under given conditions (e.g. mode of operation, stress levels, and environment). It can be quantified using metrics such as **mean time**

between failures (MTBF) or the probability of no failure within a specified period of time. MTBF is 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 without any failure for one month on the average before a failure makes the communication service unavailable. Note that the failures shorter than the survival time remain unnoticed by the application. This KPI is an end-to-end reliability metric comprising the reliability of several sub-segments, such as the equipment (including hardware and software) reliability as well as the network packet transmission reliability.

- **Resilience** can be defined as “the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions”.

Note that some of the components of dependability relate to the application perspective, while others relate to the network perspective. The relationship and mapping between these views is explained in the Appendix.

1.4 Communication Characterization

A communication service may involve different types of sources and destinations, e.g. humans and/or different types of machines such as stationary or mobile robots, controllers, actuators or sensors.

Machine-to-Machine (M2M) communication stands for the predominantly automated exchange of information between technical equipment such as machines, vending machines, vehicles or measuring systems (e.g. electricity, gas and water meters) or to a central data processing system. M2M communication includes, among other things, remote monitoring, remote control and remote maintenance of machines, equipment and systems, traditionally referred to as telemetry. Communication can be either wired or wireless. A human being is usually not involved in the communication, although a limited human involvement does not prevent the classification as M2M communication. M2M technology combines information and communication technology. Automated communication between technical equipment is predicted to be a major growth driver in the telecommunications industry in the coming years. Growth rates are expected to be many times higher than those of voice communication. The number of possible M2M devices and M2M services offered will also increase sharply in the next few years according to current expectations [BUN].

The nature of the communicating entities affects the geographical characteristics and the overall shape of the resulting traffic.

1.4.1 Geographical characteristics

The **communication area** is defined as the given area where a communication service should operate, under given conditions and requirements.

The **communication density** is the number of devices performing a given communication service per area unit, within the communication area. It may be specified in terms of minimum, maximum or average. The communication area, range and density may also be defined per volume, depending on the considered use case.

Localization/positioning error (typically expressed in m) is defined as the value of the difference between the estimated position of an object with respect to its real location, according to a reference coordinate system (which is specific for each country). Finally, knowing device **mobility aspects** is important in the design of a mobile radio network to model, among others, the handover probability, the load variation across cells or the receive power variation in both uplink and downlink [5GACIA]. Some devices might be restricted to move in predefined corridors, as this is most commonly the case for AGVs, while humans with connected augmented reality (AR) glasses are expected to move with less restriction. Therefore, the description of mobility aspects includes parameters related to the device size, velocity, motion capabilities and trajectory (predefined or not, restricted or not, pause time, etc.) as well as average distance between mobile devices. Ideally, a statistical motion model is provided (for example, the distance between vehicles is modelled as an exponential random variable with the average equal to the average distance between mobile devices).

NOTE: The moving and/or rotating sensors/actuators in a motion control system of a machine are also excluded from the mobility modelling, since typically they have little impact on the handover probability, load variation across cells, or the receive power variation [5GACIA].

1.4.2 Traffic models

Within the context of this document, traffic is said to be **symmetrical** between two nodes A and B if the average date rate from A to B is equal to that from B to A.

Communication in industrial automation has two major characteristics: **periodicity** and **determinism**.

Periodicity refers to transmitting messages with a certain transmission interval in-between messages. If that transmission interval is known and repeated, then the traffic is said to be **periodic**, otherwise it is **aperiodic**.

Determinism refers to the boundedness of the time between the transmission and the reception of a message. If this time is bounded, then the communication is said to be **deterministic**, otherwise **non-deterministic**.

Three different traffic models may be used for industrial communication, each with distinct constraints and requirements.

First, **periodic deterministic communication** is periodic with stringent requirements on timeliness and availability of the communication service. Applications producing this traffic pattern are sending messages at a fixed time interval, the transfer interval [TS22.104].

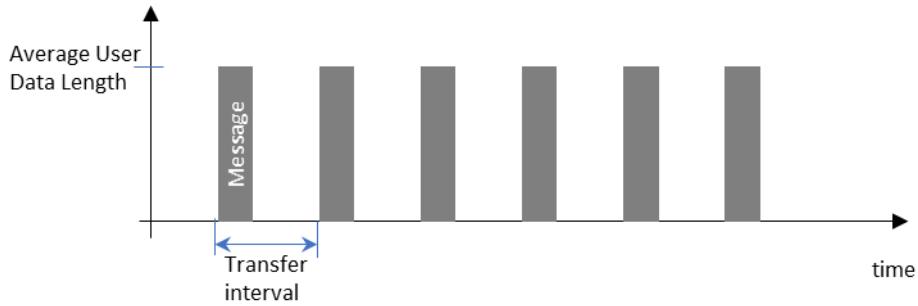


Figure 5. Periodic deterministic communications and the transfer interval [5GACIA]

The **transfer interval**, as depicted in Figure 5, is the time difference between two consecutive transfers of application data from an application to the 5G system via the CSI. Deterministic periodic traffic can be specified using the attributes of user data length (i.e., application payload size) and transfer interval [5GACIA].

Second, **aperiodic deterministic communication** consists of messages that are sent regularly but aperiodically, i.e., there is no fixed transfer interval between messages. Even without a pre-set sending time, requirements on timeliness and availability of the communication service are still stringent [TS22.104]. Applications consuming this type of traffic are still expecting to receive messages within a predictable latency [5GACIA].

Finally, **non-deterministic communication** subsumes all other traffic types than periodic/aperiodic deterministic communication. This includes periodic/aperiodic non-real-time traffic [TS22.104].

Bursty traffic is characterized by a sequence of successive messages that are sent in a burst, for example to transmit images. Bursts can occur periodically or aperiodically. It is not expected that the messages of a burst are received at the target end-point within a specified time frame [5GACIA].

1.5 Production/Manufacturing related terms

A **digital twin** is a complete and operational virtual representation of an asset, subsystem or system, combining digital aspects of how the equipment is built (PLM data, design models, manufacturing data) with real-time aspects of how it is operated and maintained [GAN18].

A **control loop** can be defined as all functional components used for automatically regulating the actual state of a process that is being controlled to match a desired state.

2 References

[NGMN15] 5G White Paper, NGMN Alliance, March 2015.

[TS22.261] 3GPP TS 22.261 V16.6.0 (2018-12), “Service requirements for the 5G system; Stage 1 (Release 16)”.

[TS22.104] 3GPP TS 22.104 V16.1.0 (2019-03), “Service requirements for cyber-physical control applications in vertical domains; Stage 1 (Release 16)”.

[NGMN2019] Verticals URLLC Use Cases and Requirements, NGMN Alliance, July 2019

[5GACIA] 5G-ACIA White Paper – Traffic Model (to be published).

[ETSI5GMEC] ETSI White Paper No. 28 “MEC in 5G Networks”, First Edition, June 2018.

[TR22.804] 3GPP TS 22.804, “Study on Communication for Automation in Vertical domains (CAV); (Release 15)”.

[IITTSN18] Time sensitive networks for flexible manufacturing testbed – description of converged traffic types, Industrial Internet Consortium, 2018.

[Gillber1960] E. N. Gilbert, “Capacity of a Burst-Noise Channel,” Bell System Technical Journal, vol. 39, no. 5, pp. 1253–1265, 1960.

[Elliot1963] E. O. Elliott, “Estimates of Error Rates for Codes on Burst-Noise Channels,” Bell System Technical Journal, vol. 42, no. 5, pp. 1977–1997, 1963

[SKTW20] P. Suskovics, B. Kovács, S. Terrill, P. Wörndl, “Creating the next-generation edge-cloud ecosystem,” Ericsson Technology Review, February 2020, <https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/next-generation-cloud-edge-ecosystems>

[ONAP] Open Network Automation Platform, <https://www.onap.org/>

[ONAP-WP] Ericsson white paper, ”ONAP and the telecom industry’s open-source journey,” March 2019, available at <https://www.ericsson.com/en/reports-and-papers/white-papers/onap-and-the-telecom-industry-s-open-source-journey>

[BSK18] Christer Boberg, Małgorzata Svensson, Benedek Kovács, ”Distributed cloud – a key enabler of automotive and industry 4.0 use cases,” Ericsson Technology Review, November 2018, available at <https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/distributed-cloud>

[GAN18] C. Ganz, “Digital twin – virtually identical?”, ABB Review, July 2018
<https://new.abb.com/news/detail/5080/digital-twin-virtually-identical>



[BUN] Bundesnetzagentur (BNetzA)

https://www.bundesnetzagentur.de/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Nummerierung/Rufnummern/M2M/M2M_node.html (in German)

3 Appendix

3.1 Derivation of communication service availability and reliability from network performance metrics

This section presents a mapping function between communication service availability and reliability requirements and network performance metrics for periodic traffic.

Throughout this section the term “packet loss” or “packet failure” is used to refer to an event in which a protocol data unit (PDU), e.g., an IP packet containing sensor updates, is not successfully delivered within a specified deadline to the target PDU layer (e.g., UPF). It is assumed that an application-level message fits in one packet on the network level.

3.1.1 Additional definitions

This clause adds supplementary terms to subsection 1.3 for the derivation of a mapping function between communication service performance and requirements on the one hand, and network level performance and requirements on the other hand.

Packet Delay Budget (PDB) defines an upper bound for the time that a packet may be delayed between the UE and the UPF¹ (bound on one-way latency) that terminates the N6 interface. For a certain 5QI² the value of the PDB is the same in UL and DL. In the case of 3GPP access, the PDB is used to support the configuration of scheduling and link layer functions (e.g. the setting of scheduling priority weights and HARQ³ target operating points).

Packet Error Ratio (PER) is the ratio of PDUs⁴ (e.g. IP packets) that have been processed by the sender of a link layer protocol (e.g. RLC⁵ in RAN of a 3GPP access) but are not successfully delivered by the corresponding receiver to the upper layer (e.g. PDCP⁶ in RAN of a 3GPP access) within the Packet Delay Budget (PDB).

Network mean time between failures (MTBF_N) is the mean time between packet failure events which represents the mean value of how long the network is available before it becomes unavailable (on a per-packet basis). Consecutive packet failures are counted as one packet failure event.

¹ User Plane Function (UPF)

² 5G Quality Indicator (5QI)

³ Hybrid Automatic Repeat Request (HARQ)

⁴ Protocol Data Unit (PDU)

⁵ Radio Link Control (RLC)

⁶ Packet Data Convergence Protocol (PDCP)

Survival time (as described in subsubsection 1.3.3) is the time that an application consuming a communication service may continue without the successful reception of an expected packet by the receiving end of the application. For periodic traffic, survival time can be expressed as maximum number of lost packets (denoted here as N_{sv} , where $N_{sv} = \left\lfloor \frac{\text{Survival Time}}{\text{Transfer Interval}} \right\rfloor$).

Network packet transmission availability is the percentage value of the amount of time the network is able to deliver packets within the agreed delay budget, 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 that the network is considered unavailable after the first packet loss. Assuming periodic traffic, it can be shown that:

$$\text{Network availability} = 1 - p = 1 - E[\text{PER}]$$

where p is the packet error probability which is the expected value of packet error ratio (PER).

Network mean time to repair (MTTR_N) is the mean value of how long the network is unavailable (on a per-packet basis) before it becomes available again. It can be derived from p and MTBF_N as:

$$\text{MTTR}_N = \frac{\text{MTBF}_N \times p}{1 - p}$$

3.1.2 Mapping function between communication service and network

To achieve the objective, and assuming periodic traffic, a Markov chain is applied, which extends the original Gilbert-Elliott Markov model [Gilbert1960, Elliot1963] to keep track of burst errors. Figure 6 illustrates, the space state is partitioned into $N_{sv} + 2$ states. The first state, U_N , represents the time that network is available on per-packet basis, i.e. it delivers packets as expected. While the network is available, the first failure happens with the probability of $MTBF_N^{-1}$. The N_{sv} middle states keep track of N_{sv} consecutive failed packets during which the communication service is still available since it is within the specified survival time. It is assumed that transition probability between the N_{sv} middle states is constant. The far-right state, D , represents the time that communication service becomes unavailable which occurs after $N_{sv} + 1$ consecutive packet failure events.

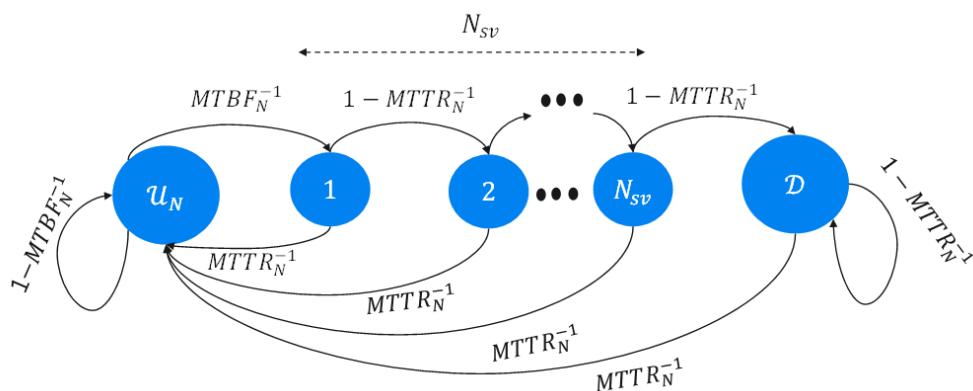


Figure 6. Markov chain for representation of burst error length

Since the proposed Markov chain is irreducible and aperiodic, it has a unique equilibrium distribution which can be derived by Markov properties using transition probabilities shown in Figure 6. In this case, the steady state probability of \mathcal{U}_N and \mathcal{D} represent the (per packet) network availability and communication service unavailability, respectively. In this case, the mean number of transitions per time unit to state \mathcal{D} can be derived as

$$\text{mean number of transitions to } \mathcal{D} \text{ per time unit} = \Pr(N_{sv}) \times (1 - MTTR_N^{-1}),$$

where $\Pr(N_{sv})$ denotes the steady-state probability of state N_{sv} . Accordingly, one over mean number of transitions per time unit is the mean time to have two consecutive transitions to state D , and therefore, multiplying this mean time to the communication service level availability results in the mean time period during which application is available. Hence, the reliability on communication service level can be calculated as

$$\text{Communication service reliability} = \frac{1 - \Pr(\mathcal{D})}{\text{mean number of transitions to } \mathcal{D} \text{ per time unit}},$$

where $\Pr(\mathcal{D})$ is the steady-state probability of state \mathcal{D} .

3.1.3 Network vs Communication service performance

Figure 7 and Figure 8 illustrate examples of the mapping between network parameters and communication service requirements when N_{sv} is 3 and 1 respectively. In Figure 7.a and Figure 8.a, the unavailability (or 1 - availability) of the communication service is shown based on network level packet error rate and mean time to repair (MTTR_N). The communication service availability requirements of 4 nines, 5 nines, 6 nines, and 9 nines are also drawn as horizontal lines. Note that the network mean time between failures (MTBF_N) can also be calculated based on PER and MTTR_N. The reliability of the communication service, which is derived based on network level MTTR_N and PER, is presented in Figure 7.b and Figure 8.b (the Transfer Interval is assumed to be equal to 1ms). The communication service reliability requirements of 1 week, 1 month, 1 year, and 10 years are also shown. It is observed that higher reliability and availability requirements put tighter requirements on the network level MTTR_N. For instance, for $N_{sv} = 3$ and PER of 10^{-6} , the average number of consecutive packet failures (i.e. MTTR_N) should be lower than 1.015 to be able to fulfil 10 years reliability requirement (refer to Figure 7.b.). However, if the application survival time is only 1 cycle ($N_{sv} = 1$), a communication service reliability of 10 years would require a network PER lower than 10^{-8} , as can be observed from Figure 8.b.

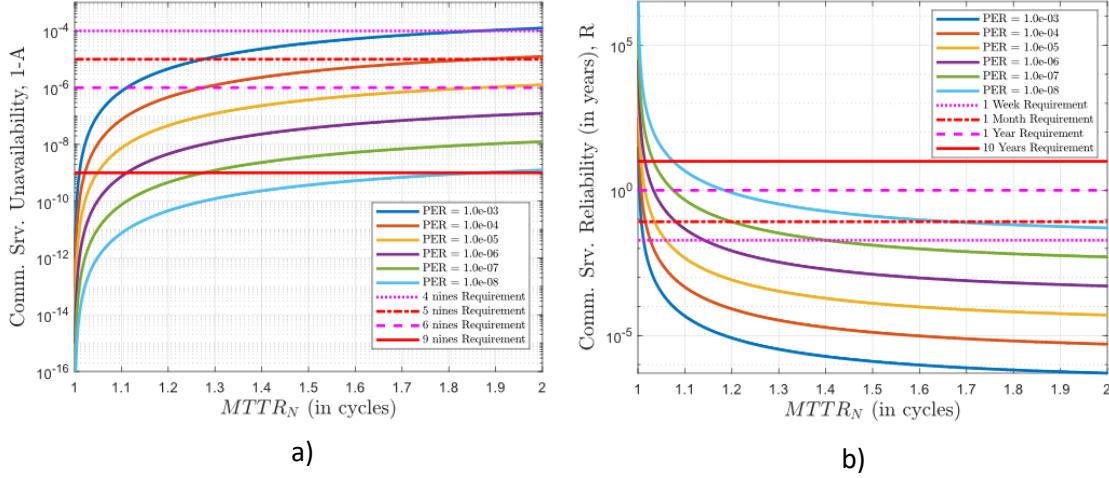


Figure 7. Impact of network level parameters on communication service availability and reliability when $N_{sv} = 3$.

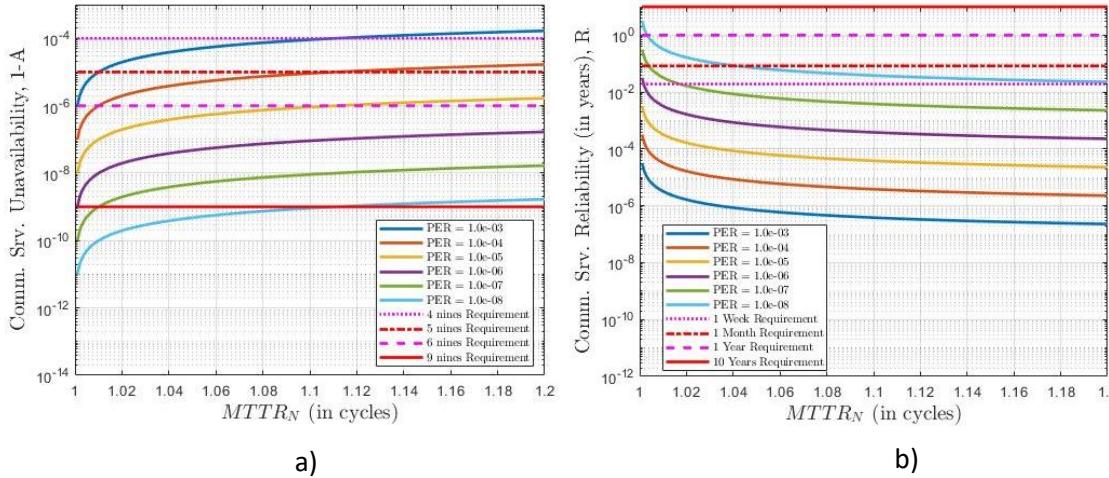


Figure 8. Impact of network level parameters on communication service availability and reliability when $N_{sv} = 1$.

List of abbreviations

AE	Acoustic Emission
AR	Augmented Reality
AGV	Automated Guided Vehicle
BBU	Baseband Unit
C2C	Controller to Controller
CNC	Computerized Numerical Control
CSI	Communication Service Interface
DL	Downlink
eMBB	enhanced Mobile Broadband
FPGA	Field Programmable Gate Array
gNB	gNodeB
GPS	Global Positioning System
HMI	Human-Machine Interfaces
I-LAN	Industrial Local Area Network
I/O	Input/Output
LTE	Long Term Evolution (3GPP technology)
MEC	Multi-access Edge Computing
mMTC	massive Machine-Type Communications
MSP	Multi-Sensor Platform
MTBF	Mean Time Between Failures
MTTF	Mean Time To Failure

MTTR	Mean Time To Repair
NFV	Network Function Virtualization
NR	New Radio (5G radio interface)
NSA	Non-Standalone (5G)
OPC UA	Open Platform Communications - Unified Architecture
OEE	Overall Equipment Effectiveness
PER	Packet Error Rate
PCB	Printed Circuit Board
PLC	Programmable Logic Controller
QoS	Quality of Service
RAN	Radio Access Network
RRH	Radio Remote Head
SA	Stand-alone (5G)
SDN	Software-Defined Networking
TCO	Total Cost of Ownership
TSN	Time-Sensitive Networking
UDP	User Datagram Protocol
UE	User Equipment
UHD	Ultra-High Definition
UL	Uplink
UPF	User Plane Function
UTM	Universal Transverse Mercator



URLLC	Ultra-Reliable Low Latency Communications
VR	Virtual Reality