http://dx.doi.org/10.25673/39578

Co-configuration of 5G and TSN enabling end-to-end quality of service in industrial communications

Lukas Martenvormfelde[1], Arne Neumann[1], Lukasz Wisniewski[1], and Lukas Schreckenberg[2]

[1] inIT - Institute industrial IT, Technische Hochschule Ostwestfalen-Lippe, Campusallee 6, 32657 Lemgo, Germany
{lukas.martenvormfelde, arne.neumann, lukasz.wisniewski}@th-owl.de
[2] Fraunhofer IOSB-INA, Campusallee 1, 32657 Lemgo, Germany
lukas.schreckenberg@iosb-ina.fraunhofer.de

Abstract. The recent trends in automation require a highly reliable communication in order to enable distributed controls in real-time applications. Time Sensitive Networking and 5G promise to provide the required Quality of Service for the wired and wireless communication, respectively. However, high reliability comes at the expense of a high configuration and engineering effort especially for heterogeneous networks. This paper aims to reduce the barriers of a successful co-configuration 5G and TSN. Therefore, a broad overview over the specifications of both technologies is given with a focus on the individual and joint configuration of both systems. Subsequently, architectural aspects of the co-configuration and a mapping of 5G and TSN parameters are pointed out to enable end-to-end quality of service in industrial communications.

1 Introduction

In recent years, the changes towards Industry 4.0 increased the demands on the underlying communication infrastructure. Particularly, it is required to reliably connect hundreds and thousands of devices some of which have stringent requirements on data rates, error rates or timeliness of the packets. Since it would hardly be possible to cover all aspects in a single communication technology, industrial communication is and will be characterized by combining different network technologies bringing all their beneficial sets of features into a network [16]. The efficient configuration, deployment and maintenance of the resulting heterogeneous networks is considered as a relevant challenge in the area of industrial networks [6]. With respect to demanding applications, several network parameters have to be thoroughly adjusted in order to achieve the required quality of service. Consequently, combining two or more network technologies together is a not straight forward task.

5G and Time Sensitive Networking (TSN) are two communication technologies that gained a lot of interest in industrial communication systems over the last few years. TSN represents a wired Ethernet based technology defined by the IEEE 802.1 TSN task group providing a set of standards for precise clock synchronization and real-time traffic treatment which is seen as an enabler for time-critical communications [11, 12]. If mobility requirements apply, reliable wireless communication is required in complement to the wired communication technologies, and due to the scarcity of radio spectrum, cognitive radio extensions to unlicensed wireless technologies [3] or the use of licensed systems such as 5G is essential. With 5G, the 3GPP targets enhanced Mobile Broadband (eMBB) as well as massive Machine Type Communication (mMTC) and Ultra Reliable Low Latency Communication (URLLC) in order to fulfil the manifold demands of various applications [7]. Moreover, an integration approach of modelling the 5G system as a bridge in the TSN network has been introduced in research articles [13, 14] and specified by the 3GPP, emphasizing the function of the TSN Translator at User Equipment side and Core Network side of the 5G system.

This paper gives a brief overview over TSN stream types which are defined in IEC/IEEE 60802 and its 5G counterpart, the Quality of Service (QoS) flows, in Section 2. In contrast to many papers that analyse the delays and reliabilities of TSN over 5G [10] or the individual technologies, this paper focuses on aspects of the joint configuration of streams or flows in 5G/TSN network carried out in Section 3. Finally, the work is concluded in Section 4.

2 State of the Art

2.1 Quality of Service Classes in conjunction with Ethernet TSN and usage of TSN mechanisms of the TSN profile 60802

TSN enhances classical Ethernet with mechanisms for clock synchronization and real-time traffic handling by providing a set of several functions. Unfortunately, different implementations and standards use different variations of this functions. For example CC-Link IE TSN uses synchronization and Time Aware Shaper (TAS), but does not use preemption. PROFINET defines synchronization, preemption and TAS for 100 Mbps. In this paper the focus is put on the IEC/IEEE 60802 TSN Profile for Industrial Automation, draft 1.3. [8]. The real-time traffic consists of time-sensitive streams that are defined as a unidirectional stream of data frames which have to be delivered within a bounded time. Streams and network devices are managed in Ethernet TSN Domains. [9]

Due to IEC/IEEE 60802 still being in a draft status, the PROFINET over TSN Guideline, which attempts to anticipate IEC/IEEE 60802, is used as a reference [5]. Three different stream types are defined, which are depicted in table 1 together with common best effort traffic for comparison. The *high stream* is used for cyclic isochronous traffic where the network interface and the application are synchronized. For this stream type, a TSN end station knows when to sent a certain frame. This is called a synchronized network access. Moreover, the routes for the communication are engineered by a central Network Management Engine (NME), also called TSN Domain Management Entity (TDME) and Media Access Control (MAC) address learning is disabled. In contrast the *low stream* type is used for non-isochronous traffic where the network interface and the application are not synchronized. For the real-time stream type there is no time awareness, but it is still cyclic traffic which has a latency requirement. Furthermore the routes are not engineered but learned through the Spanning Tree algorithm. [5]

Table 1. TSN stream types and best effort traffic adapted from PROFINET over TSN Guideline Version 1.31

Stream Type	**Cyclic /Acyclic**	**Latency Requirement**	**MAC Learning /Engineered**	**Time Awareness**
High (isochronous)	cyclic	yes	engineered	yes
Low (non-isochronous)	cyclic	yes	engineered	yes
Real-Time	cyclic	yes	learning	no
Best Effort (no stream shown for comparison)	acyclic	no	learning	no

The mapping of traffic to a specific stream and stream type is defined in the Ethernet frame header. To realise that, both the destination address and the Virtual Local Area Network (VLAN)-Tag of the frame are used. The tag includes the Priority Code Point (PCP) field in which the stream type is coded as a priority number as well as the VLAN Identifier (VID) field. The VID in combination with the destination address is used as a unique stream path selector. According to IEC/IEEE 60802 draft 1.3 there shall be a middleware with a translation table to translate between different priority types for applications not using the specified PCP or VID. This middleware can also translate to other profiles like 5G [8].

As described in the PROFINET over TSN Guideline [5], there are several mechanisms to ensure the determinism and low latency of the traffic. In order to protect a TSN domain against too much incoming external traffic, e.g. from a different TSN domain, which could block the internal streams, ingress rate limiters can be used. These function of Ethernet TSN bridges enforce a bandwidth limit for the external traffic. Moreover, a new priority and VID is assigned to the incoming traffic when entering the new TSN domain. As the traffic leaves the domain, the previously assigned priority and VID are removed.

Further interdependent mechanisms influence the traffic handling at bridges. When leaving a bridge, the streams are inserted in different queues depending on the assigned priority and are sent only when the permission for that queue is granted by a Transmission Selection Algorithm (TSA). Using only the priority queues is the most basic algorithm which is called strict-priority. However, it can be enhanced with the TAS which adds a time schedule for the different queues and requires common notion of time among the TSN domain members. For instance, the permission for the queue with *high streams* can be always given, while the permission for the *low stream* is only given for a shorter time slot of the communication cycle and the permission for the real-time streams is given for the shortest time slot. The configuration of time aware shaping in a TSN can be generalised e.g. as a no-wait job shop scheduling problem which has been proven to be NP-hard [4]. Therefore, typically some heuristic approaches are used to schedule such communication. To prevent enlarged latencies of *high streams*, through low priority traffic extending into their time slots, guard bands are used. A guard band is a gap in the transmission as large as the largest possible Ethernet frame. This gap, serving as a buffer, is placed in front of the scheduled time slot of the *high stream*. To utilize the limited available time slots for transmissions more efficiently preemption is introduced in TSN. It is the procedure of splitting a low priority frame into parts and sending them separately. Because of preemption the guard band can be minimized to a size as small as the smallest possible Ethernet frame. Forwarding techniques reduce the latency when a transmission is send over a bridge. Usually the bridge would buffer the whole frame and forwards is afterwards. With the cut-through forwarding method, the incoming frame is directly forwarded to the output port without buffering it. Still a short delay due to the processing in the hardware is existent. In IEC/IEEE 60802 draft 1.3 a delay less than 1 µs is recommended and less than 2 µs is mandatory for a bridge with a specified bandwidth of 1 GB. Cut-through forwarding is only possible, if the port is not blocked by an other transmission.

2.2 5G QoS Flows

The 5G system architecture as depicted in Fig. 1 is described in TS 23.501 [2]. A 5G system is split into a user plane which is visualized by the larger boxes on the bottom of the illustration and a control plane depicted on top of it. A typical user data flow goes from the Data Network (DN), i.e. some external application, through the User Plane Function (UPF) and the Radio Access Network (RAN) to a User Equipment (UE) or vice versa. The control plane consists of 22 different network functions some of which may have multiple instances. A list of all instances of network functions is held in the Network Repository Function (NRF). Two of the most notable and irreplaceable control plane functions are the Access and Mobility Function (AMF) and Session Management Function (SMF) which connect the user plane to the control plane and thus one of their key responsibilities is to pass configuration parameters to the UE, RAN, and UPF, respectively. Furthermore, the Application Functions (AFs) which represent the application interface to the 5G control plane are as dynamic as the related application itself and may start or stop depending on the Operation Technology (OT) requirements. If an AF is trusted by the 5G core, the application can influence parts of the 5G network behavior such as the traffic routing. Moreover, a Network Exposure Function (NEF) exposes network capabilities and statistics to external networks and applications. Thus, an interaction between an OT application and the 5G control plane can be realized over the NEF and the AF.

Furthermore, the 3GPP describes several aspects related to the QoS of the 5G system as well as the integration of 5G with TSN [2]. The specification describes the flow-based QoS model as the finest granularity to differ between the packet flows. Each flow is identified by a unique QoS Flow ID (QFI), and within a PDU session at least one default flow needs to be established. Further QoS

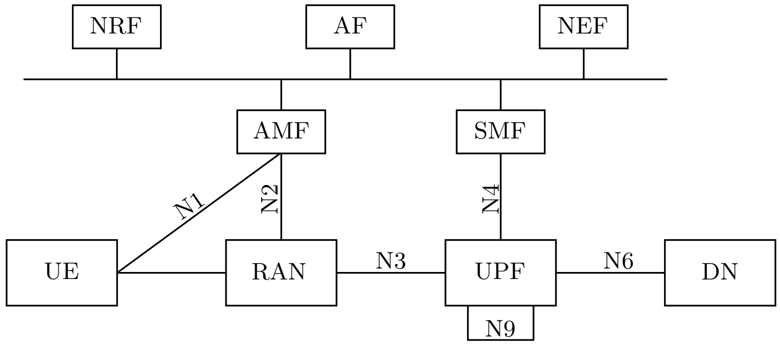

Fig. 1. 5G system architecture

flows may be configured during the PDU session establishment or dynamically assigned by the SMF of the 5G system. Packets that arrive at the 5G system require a mapping to the corresponding QoS flow and finally the radio resources as illustrated in Fig. 2. Therefore, packet detection rules are applied to filter the packets in the UPF or the UE, respectively. The filtering of the packets is supported based on IP or Ethernet headers including the IEEE 802.1Q [9] fields, and the packets mapped to a QoS flow of a PDU session are then forwarded to the RAN with an additional QFI encapsulation. The QFI is used by the Service Data Adaption Protocol (SDAP) layer in the RAN to detect the QoS profile and thus enables the mapping of the flows to the available radio resources, i.e. the Data Radio Bearers (DRBs), but the QFI is not transmitted over the radio interface.

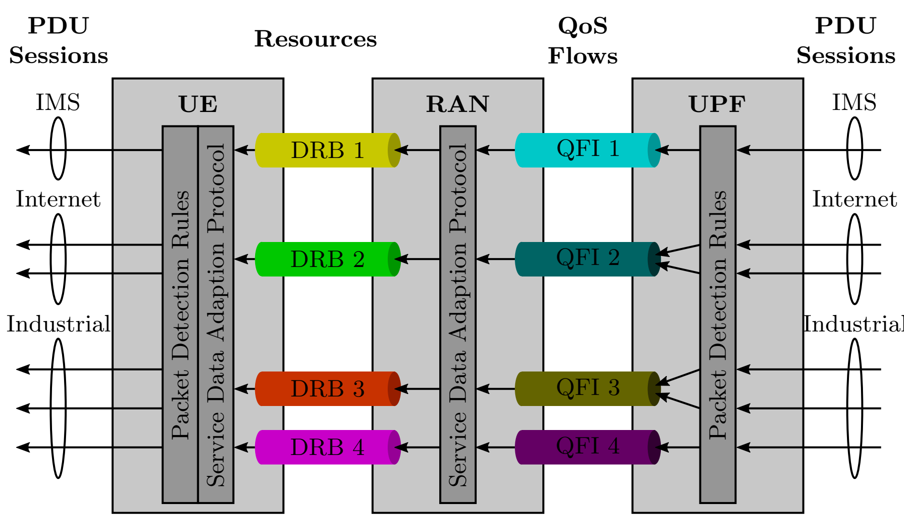

Fig. 2. QoS Flow Model

The flows can be parametrized by the 5G QoS Identifier (5QI) value that refers to various QoS characteristics and has predefined QoS mappings specified by the 3GPP. Furthermore, an Allocation and Retention Priority can be set which is examined by the RAN in order to decide which flows can be preempted in case of resource constraints. In 3GPP networks, emergency flows are assigned to the highest priority and thus they may not be preempted but any other flow may be preempted in order to enable the emergency tasks. Typical preemptable flows concern the user traffic in internet PDU sessions. For industrial purposes, the Allocation and Retention Priority might range somewhere in between and might also be used to distinguish between the importance of the different factory applications. Moreover, the Guaranteed Flow Bit Rate (GFBR) and Maximum Flow Bit Rate (MFBR) of a flow can be set with respect to the required minimal bit rate and the expected maximal bit rate of the initiated flow. Furthermore, the aggregated maximum bit rates for a UE and a PDU session and a maximum tolerable packet loss rate, specified for a QoS flow can be set. The QoS characteristics related to the parametrized 5QI value contain the resource type. This can be of type Guaranteed Bit Rate (GBR) or non-GBR as well as a priority level affecting resource scheduling among different flows and UEs. Contrary to the Allocation and Retention Priority, the priority level of the QoS profile does not abort the entire flow but may prioritize one flow over another. Additionally, a Packet Delay Budget is given to define a maximum delay that the packet forwarding in the 5G system may require. Also a maximum Packet Error Rate for the non-congested system is used in order to provoke changes on the lower layer properties such as the code rate or the modulation scheme. For GBR flows, the averaging window can be specified over which the bit rate shall be guaranteed. If the GBR flow is delay critical, the Maximum Data Burst Volume (MDBV) needs to be provided. All the above mentioned configurations are done by the SMF by signaling either the 5QI value of a predefined QoS profile or each of the parameters.

2.3 5G integration with TSN

With Release 16, the 3rd Generation Partnership Project (3GPP) highlights the integration of 5G in time sensitive communications and defines an architecture for 5G as a TSN bridge in TS 23.501 [2]. The specification is not limited to a specific deployment such as Ethernet TSN. For TSN over 5G, the 5G system is encapsulated, by two TSN-Translator functions on the device side (DS-TT) and network side (NW-TT) in order to enable the 5G integration as shown in Fig. 3. The term TSN system as used in the illustration refers to any TSN network beginning from the smallest possible unit, i.e. a single TSN host. Only one PDU session between the DS-TT and the UPF can be established for a 5G TSN bridge but due to the N9 interface that can connect multiple UPFs, a UE with multiple DS-TTs can be part of multiple logical TSN bridges. The 5G QoS flows for TSN integration are delay-ciritcal GBR flows.

Inside the 5G control plane, a TSN application function is used to connect with the TSN network controller. As mentioned previously, the AF can affect the

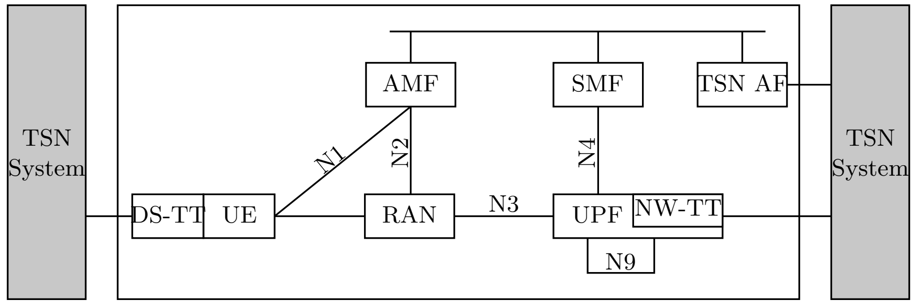

Fig. 3. 5G Bridge Architecture

configuration of the 5G network. In order to obtain the behavior of a transparent bridge as suggested in [10, 13–15], the TSN AF needs to hold information about the delays inside the 5G system and needs to provide forwarding rules to forward layer 2 Ethernet TSN frames over the 5G system operating on layer 3. The standard mainly deals with time synchronization aspects of the 5G integration and specifies that the 5G system appears as a slave to the master clock of the TSN system. Incoming packets get an ingress timestamp by the TSN translator function, and an egress timestamp is created before a packet leaves the 5G system. Moreover, the translator functions shall support hold & forward buffering mechanisms to reduce the jitter of deterministic TSN flows.

TS 23.501 [2] and TS 23.503 [1] further specify details about the management and configuration of the 5G system as a TSN bridge. The interface between the TSN AF and the Centralized Network Configuration (CNC) of the TSN system is used to exchange information about the TSN configuration. For instance, a-priori information about the traffic characteristics such as the periodicity or the frame sizes can be passed from the CNC to the TSN AF. If the characteristics are unknown, the TSN AF can calculate traffic patterns. However, in both cases the TSN AF is required to send Time Sensitive Communication Assistance Information (TSCAI) to the SMF in order to configure the 5G system. In particular, the RAN benefits from the TSCAI which allow semi-persistent scheduling or configured grant transmissions.

3 Co-configuration aspects

Some architectural requirements for self-configuration of the network which are not finally standardized yet or out of scope of the standardization emerge from the specification of the above described technologies. With respect to the joint co-configuration of 5G and TSN, particularly, the connection between the 5G control plane in form of the TSN AF and the CNC of the TSN system is of importance. Since the TSN system is the superior system in the architecture, the CNC is expected to initiate the configuration. However, a fully automatic configuration requires further development and thus the following discussion will highlight parameters that are of interest rather than the automatic exchange of information. At first, the physical layer configuration of the 5G system has to be considered. Although it is out of scope of this paper, it has to be mentioned that the bandwidth of the 5G system, the duplexing mode and the numerology, e.g. the granularity in time and frequency, are important for the QoS of deterministic traffic and may be tweaked with respect to the desired communication type if accessible. Besides the physical layer properties of 5G which set the boundaries for the communication, many other aspects need to be taken into account in order to map the 5G QoS flows to the TSN streams.

Table 2. Parameter mapping of 5G system to TSN stream

TSN	**5G**	**Description**
Frame Size	GFBR	QoS flow guarantees the resources for the TSN frame
Frame Size	MDBV	No more data than the usual TSN frame has to be transmitted
Periodicity	Averaging Window	Periodicity of the TSN stream is identical to the Averaging Window in which the GBR of the QoS flow is measured

A TSN stream can be mapped to a 5G QoS flow by applying the related Packet Detection Rule at the UPF or UE which can be based on the source and destination addresses as well as the PCP and VID. However, the SDAP mapping of the QFI to the DRB and the scheduling might still lead to suboptimal transmission slots. Obviously, the QoS flow in the 5G system needs to be of type GBR unless it deals with a best effort TSN stream class. Assuming a cyclic TSN stream with constantly large frames or bursts of frames, the 5G QoS flow further needs to guarantee the GFBR within an averaging window which equals the periodicity of the TSN stream. Moreover, since bursts of one or more packets are expected, the MDBV needs to be as large as the amount of data which has to be transmitted within the averaging window as summarized in Table 2. But what will the 5G system make out of this configuration? The GFBR guarantees that enough resources on the DRB are allocated for the transmission of the flow, and the MDBV causes a compact allocation of the resources instead of spread time slots. Thus, all packets or packet fragments of the flow will be transmitted nearly at the same time, but however, this mechanism still leaves plenty of space for decisions to the scheduler of the 5G system which presumably introduces an increased jitter. Nonetheless, the jitter can be reduced by the buffer & hold mechanisms of the 5G TSN translator functions which leads to a trade-off between jitter on the one hand and delay on the other hand with both being undesirable for the TSN stream types *high, low, or real-time*. An easy but inefficient approach might be to reserve more resources than actually required by increasing the GFBR or fractioning the averaging window at the expense of blocked resources. Moreover, semi-persistent scheduling or configured grants can be exploited in order to reduce the jitter. However, the complexity of the scheduling which is already an NP-hard optimization problem for the scheduled TSN traffic further increases with this approach. Nevertheless, a 5G system will likely reduce the total amount of switches between two TSN end stations.

4 Conclusion

This paper summarized essential QoS related mechanisms of TSN and 5G as a state-of-the-art review of relevant specifications. Moreover, a direction for the joint configuration of 5G and TSN was carried out showing that a significant amount of considerations has to be taken into account. Parameters of both communication domains have been mapped to each other to provide an impression of the ongoing challenges in the joint configuration of heterogenous networks combining 5G and TSN.

In the future, the considerations of this paper have to be implemented in order to validate the concepts and collect the first results. Moreover, the high complexity of configuration and engineering of both systems needs to be reduced. Due to the diversity of different application requirements in modern as well as future factories, it is questionable if the prioritization among the different types is feasible. It has to be studied if a coarser differentiation and planning of not all but the highest priority is sufficient in order to reduce the complexity and ensure the QoS especially for the streams of the highest importance.

5 An author's note on 5G and TSN terminologies

Although both 5G and TSN are communication technologies for industrial real-time communication, the domains are separated enough to do not share a common terminology. The authors of this paper stumbled over several terms and acronyms that have a different meaning in the domains of 5G and TSN and want to share this experience in order to avoid future misunderstanding. For instance, both technologies utilize preemption, but while TSN preempts the transmission of a single Ethernet frame, 5G preempts PDU sessions and QoS flows based on the ARP which does relate to the Allocation and Retention Priority instead of the widely known Address Resolution Protocol.

References

1. 3GPP: Policy and charging control framework for the 5G System (5GS). Tech. Rep. TS 23.503, 3GPP (2021)
2. 3GPP: System architecture for the 5G System (5GS). Tech. Rep. TS 23.501, 3GPP (2021)
3. Chiwewe, T.M., Mbuya, C.F., Hancke, G.P.: Using cognitive radio for interference-resistant industrial wireless sensor networks: An overview. IEEE Transactions on Industrial Informatics **11**(6), 1466–1481 (2015). https://doi.org/10.1109/TII.2015.2491267

4. Dürr, F., Nayak, N.G.: No-wait packet scheduling for ieee time-sensitive networks (tsn). In: Proceedings of the 24th International Conference on Real-Time Networks and Systems. p. 203–212 (2016). https://doi.org/10.1145/2997465.2997494
5. Friesen, A., Schriegel, S., Biendarra, A., Gamper, S.: PROFINET over TSN Guideline Version 1.31. In: Profibus International (PI) (2021)
6. Gaj, P., Scanzio, S., Wisniewski, L.: Guest editorial: Heterogeneous industrial networks of the current and next-generation factories. IEEE Transactions on Industrial Informatics **16**(8), 5539–5542 (2020). https://doi.org/10.1109/TII.2020.2976796
7. Ghosh, A., Maeder, A., Baker, M., Chandramouli, D.: 5g evolution: A view on 5g cellular technology beyond 3gpp release 15. IEEE Access **7**, 127639–127651 (2019). https://doi.org/10.1109/ACCESS.2019.2939938
8. IEC/IEEE: Time-sensitive networking profile for industrial automation. IEC/IEEE 60802, Draft 1.3 pp. 1–104 (2021)
9. IEEE: IEEE standard for local and metropolitan area network–bridges and bridged networks. IEEE Std 802.1Q-Rev Draft 1.0 (Revision of IEEE Std 802.1Q-2018) (2021)
10. Larrañaga, A., Lucas-Estañ, M.C., Martinez, I., Val, I., Gozalvez, J.: Analysis of 5g-tsn integration to support industry 4.0. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). vol. 1, pp. 1111–1114 (2020). https://doi.org/10.1109/ETFA46521.2020.9212141
11. Leßmann, G., Biendarra, A., Schriegel, S.: Vergleich von Ethernet TSN-Nutzungskonzepten. In: 2020 10th Annual Colloquium Communication in Automation (KommA2020) (2020)
12. Lo Bello, L., Steiner, W.: A perspective on ieee time-sensitive networking for industrial communication and automation systems. Proceedings of the IEEE **107**(6), 1094–1120 (2019). https://doi.org/10.1109/JPROC.2019.2905334
13. Mannweiler, C., Gajic, B., Rost, P., Ganesan, R.S., Markwart, C., Halfmann, R., Gebert, J., Wich, A.: Reliable and deterministic mobile communications for industry 4.0: Key challenges and solutions for the integration of the 3gpp 5g system with ieee. In: Mobile Communication - Technologies and Applications; 24. ITG-Symposium. pp. 1–6 (2019)
14. Neumann, A., Wisniewski, L., Ganesan, R.S., Rost, P., Jasperneite, J.: Towards integration of industrial ethernet with 5g mobile networks. In: 2018 14th IEEE International Workshop on Factory Communication Systems (WFCS). pp. 1–4 (2018). https://doi.org/10.1109/WFCS.2018.8402373
15. Neumann, A., Wisniewski, L., Musiol, T., Mannweiler, C., Gajic, B., Ganesan, R.S., Rost, P.: Abstraction models for 5g mobile networks integration into industrial networks and their evaluation. In: Jahreskolloquium Kommunikation in der Automation - KommA 2018. Lemgo, Germany (2018)
16. Wollschlaeger, M., Sauter, T., Jasperneite, J.: The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0. IEEE Industrial Electronics Magazine **11**(1), 17–27 (2017). https://doi.org/10.1109/MIE.2017.2649104