Reliable and Deterministic Mobile Communications for Industry 4.0: Key Challenges and Solutions for the Integration of the 3GPP 5G System with IEEE Time-Sensitive Networking

Christian Mannweiler[1], Borislava Gajic[1], Peter Rost[1], Rakash S. Ganesan[1], Christian Markwart[1], Rüdiger Halfmann[1], Jens Gebert[2], and Andreas Wich[2]
[1]Nokia Bell Labs, München, Germany
[2]Nokia Bell Labs, Stuttgart, Germany
{christian.mannweiler,borislava.gajic,peter.m.rost,rakash.sivasivaganesan,christian.markwart,ruediger.halfmann, jens.gebert,andreas.wich}@nokia-bell-labs.com

在工业4.0环境中，提出了中心化控制、5GS作为逻辑桥和TSN集成部署的方案，并提供了5GS和TSN之间的QoS映射框架。

Abstract

The digitization of industrial automation processes calls for flexible, adaptable, and scalable communication solutions in order to enable the vision of a truly "cyber-physical system". In this context, IEEE Time Sensitive Networking (TSN), developed by the TSN task group of the IEEE 802.1 working group, receives particular interest as it defines mechanisms for the time-sensitive (i.e., deterministic) transmission of data over Ethernet networks. In order to enable novel use cases and further improve the efficiency of industrial automation, mobile communication solutions are needed. Here, the time-sensitive communications (TSC) service in 3GPP Release 16 aims to support applications requiring deterministic or isochronous communication with high reliability and availability, such as IEEE TSN and IETF DetNet, over 5G mobile networks.

This paper briefly depicts the favorable deployment scenarios for an integrated 5G/TSN system in Industry 4.0 environments. It further describes the requirements in integrating a 5G system (5GS) with an IEEE TSN system and depicts the 3GPP integrated system architecture, which foresees to present the 5GS to the TSN system like any other TSN-aware bridge. Subsequently, the focus of this paper lies on presenting remaining challenges as well as solutions to a subset of these challenges, particularly focusing on radio access, core, and network management aspects as well as QoS mapping framework.

1 Introduction

One of the critical enabler for Industry 4.0 [1] is scalable and pervasive connectivity between people, machines, and other objects. Today, the vast majority of communication technologies used in manufacturing industry is wire-bound. This includes a variety of dedicated Ethernet-based technologies (e.g., Sercos®, PROFINET®, and EtherCAT®) and field buses (e.g., PROFIBUS®, CAN®, etc.) [2, 4]. To overcome this heterogeneity (异质性), the objective of IEEE Time-Sensitive Networking (TSN) task group is to provide deterministic services through IEEE 802 networks, i.e., guaranteed packet transport with bounded latency, low packet delay variation, and low packet loss. It has defined several standards covering aspects, such as, synchronization, stream reservation, preemption, scheduling, and frame replication and elimination for reliability (FRER).

In parallel, wireless connectivity has become an important means for increasing flexibility in industrial production [3]. To date, wireless communication has primarily been used for non-time-critical applications and IT-originated scenarios. The most significant advantage of introducing wireless communication also for critical applications is the provisioning of dependable connectivity to moving objects such as mobile robots, automated guided vehicles (AGVs), drones, and humans [5]. The 3GPP 5G system (5GS) can provide such wireless communication services.

In general, there are various possible options for the 5GS to deliver a TSN service or to be integrated into an existing TSN system. Figure 1a shows a stand-alone deployment, where the 5GS directly supports a TSN service, enabling TSN end stations to be connected directly to ports at either user equipment (UE) or user plane function (UPF). In Figure 1b, it is shown how the 5GS may also be integrated as a self-contained solution into an existing TSN system. In this example, the TSN end stations using mobile connectivity are still directly connected to UEs. Alternatively, a UE could be connected to other wire-bound TSN bridges to extend the TSN system, e.g., if a mobile device (e.g. AGV) hosts several TSN end stations. Figure 1c shows how a distributed 5GS deployment may re-use the TSN system to transfer data and signaling traffic between 5GS network elements. In contrast to the scenarios in Figures 1a-1c where only one TSN domain is considered, Figure 1d shows a case where a single 5GS deployment simultaneously serves multiple TSN domains.

The remainder of this paper briefly introduces fundamentals of the 3GPP 5GS and IEEE TSN, in Section 2. Section 3 describes the requirements to integrate the two systems, depicts the integration architecture defined in 3GPP. Section 4 provides a selection of remaining challenges and a solution for Quality of Service (QoS) mapping between the 5GS and the TSN systems. Finally, Section 5 concludes the paper.

ISBN 978-3-8007-4961-4

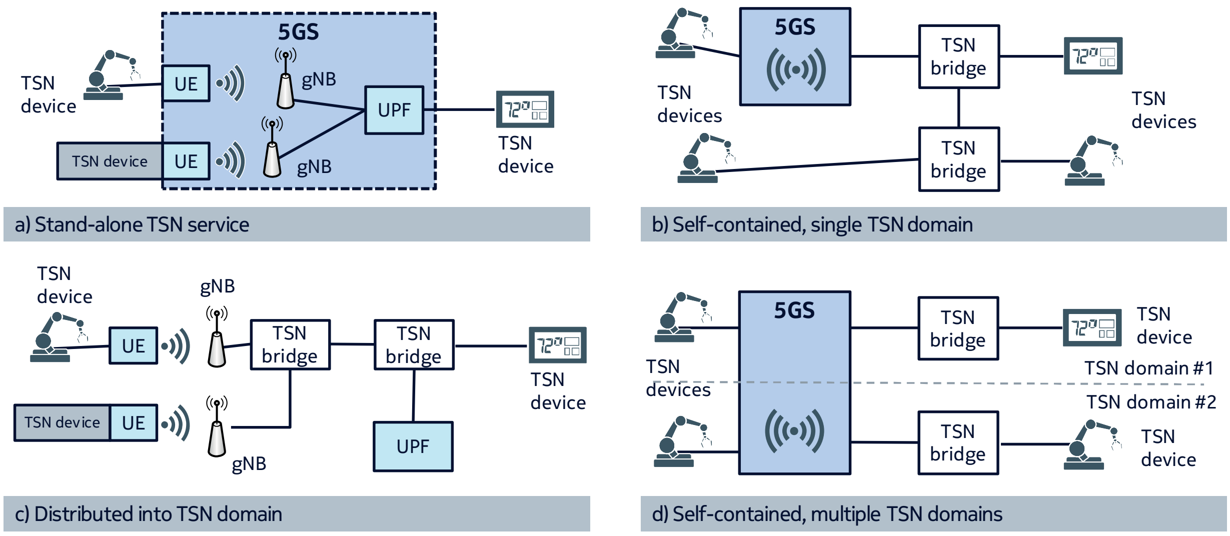

Figure 1 Scenarios for integration of 3GPP 5GS and IEEE TSN

2 Standalone systems: 3GPP 5GS and IEEE TSN

2.1 3GPP 5G System

The 5G standardization in 3GPP is divided into two phases. In the first phase, Rel. 15 addressed a more urgent subset of use cases to meet commercial needs (September 2018) and in the second phase, Release 16 will target the ITU IMT-2020 submission (December 2019). The architecture of the 3GPP 5G System (5GS) is shown in Figure 2 [10] and it consists of three essential components: the user equipment (UE), the radio access network (NG-RAN), and the 5G core network (5GC). The core network is divided into UPFs and control plane (CP) involving access and mobility management (AMF), session management (SMF), policy control (PCF), and unified data management (UDM) functions. The 3GPP 5GS is connected with non-3GPP data networks (DN) through the N6 interface and on CP with application functions (AF) through N5.

Among many other key features, the 3GPP 5GS applies a new form of defining interfaces (based on "services"), it separates user and control plane, it supports network slicing, and it can integrate different access network technologies including 3GPP LTE.

Currently, support for local area networking (LAN) and specifically the support of time sensitive communication (TSC) receives significant interest within the 3GPP, which is captured in dedicated study items describing requirements and use cases. Naturally, these use cases will have an impact on the 3GPP architecture and RAN design which is being addressed in the 3GPP study items "5GS Enhanced support of Vertical and LAN Services" [11] and "New Radio Industrial IoT" [13], respectively. These study items investigate, among others, determinism and synchronization down to sub-microseconds, integration with IEEE TSN,

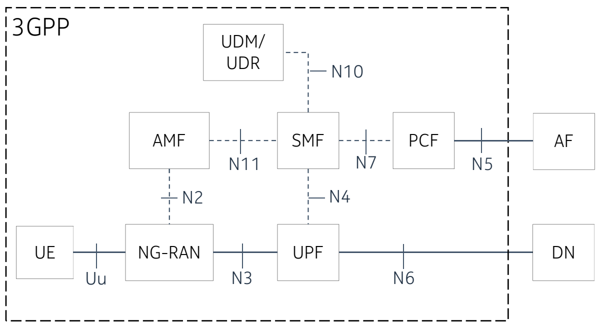

Figure 2 5G system architecture

and enabling full support for deterministic communications scaling from industrial LANs up to multi-hop and IP WAN deployments, similar to IETF DetNet [6].

2.2 IEEE Time-Sensitive Networking

TSN [8] is a collection of IEEE 802.1Q protocols [9] which, when configured appropriately, result in an E2E communication with very low and deterministic delay. TSN implements three different configuration models, i.e., fully centralized model, centralized network/distributed user model, and fully distributed configuration model. In this paper, we focus on the fully centralized model although the described architecture integration equally applies to the other models. Figure 3 shows the fully centralized configuration model of the TSN network described in IEEE 802.1Qcc.

The talker end stations are devices such as sensors providing information to listener end stations, which consume the information, e.g., controllers or monitoring devices. The network consists of a set of time-aware bridges. These bridges and the end stations are synchronized to a mas-

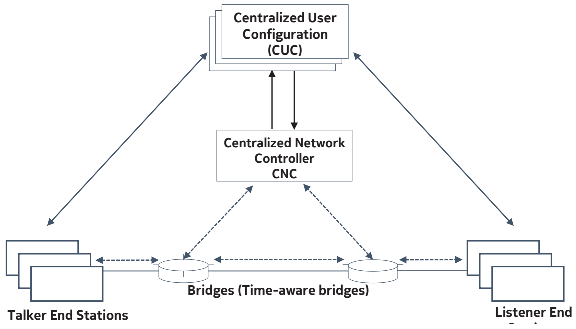

Figure 3 TSN fully centralized configuration model

	Use Case Class (UCC)		
Requirements	UCC A	UCC B	UCC B
Service availability	99,9 – 99,9999 %	99,99 – 99,999999 %	99,9999 – 99,999999 %
E2E latency	10 ms – 1 min	1 – 100 ms	0.25 – 5 ms
Synchronicity	0.5 ms – 0.5 s	10 µs – 50 ms	0.25 µs – 1 ms
Message Size	40 B – 10MB	20 – 250 B	20 – 100 B
Transfer Interval	20 ms – 1 day	1 – 500 ms	< 0,5 – 10 ms

Table 1 Industry 4.0 requirements for periodic message exchange

ter clock in the system and hence, they are aware of the global absolute time in the system. The set of all communication partners in one time domain (talkers and listeners) constitutes a TSN domain. The end stations provide their communications requirements (TSN streams) to the Centralized User Configuration (CUC), which translates the stream requirements into corresponding communication requests towards the network. In one TSN domain, multiple CUCs may co-exist. The CUCs are ecosystem-specific such as specific CUCs for PROFINET or OPC UA. Each TSN domain has a single entity called Centralized Network Controller (CNC). The CNC receives the E2E stream requests from the CUCs through the user/network configuration interface, which is standardized in TSN. A CNC has a complete view of the network in order to compute the transmission schedule and configure each of the bridges so that E2E communication can be established. In the following, we describe how the 3GPP 5GS and the IEEE TSN can be integrated in an efficient manner.

3 Integration of 3GPP 5G System and IEEE TSN

3.1 Integration requirements

In the TACNET 4.0 project, more than twenty representative use cases related to manufacturing and automation in industry environments were studied to derive their specific requirements on a wireless communication service [7]. Table 1 shows in an exemplary way the key requirements regarding the periodic message exchange for each of the identified use cases.

Use Case Class (UCC) A includes, among others, Process Automation, Additive Sensing, Remote Monitoring, and Predictive Maintenance. UCC B includes, among others, Remote Control, Discrete Automation, Mobile Robots, Closed Loop Process Automation, Safety Applications, Augmented Reality, and Cooperative Transport of Goods. UCC C includes, among others, Cooperative Robotics, Motion Control, Closed Loop Motion Control.

Beside the periodic traffic exchange, each use case also requires further aperiodic message exchange to address the needs of operation, administration, safety, and maintenance tasks. Examples for such tasks include the deployment of control logic modifications as well as software and security updates. Even if the requirements of aperiodic message exchange on service availability as well as on latency are typically more relaxed than for the periodic message exchange, it should be noted that different tasks may require task-specific QoS handling to guarantee a reliable operation of industrial applications.

3.2 Integration architecture: Using 3GPP 5GS within TSN systems

The architectural options for the inter-working of the 3GPP 5GS with a TSN system comprise two fundamental approaches. As a first option, the 5GS can be modeled as a TSN Ethernet link. In this case, the 5GS is expected to behave like a “twisted pair cable” with deterministic performance in terms of capacity and delay. This entails a couple of disadvantages and restrictions, among them the lack to adequately modeling the dynamic QoS performance inherent to a wireless system and the very limited possibility to influence the scheduling of TSN streams.

In contrast, the second option “5GS Bridge”, which models the 5GS as a logical TSN bridge [11], provides several advantages and can more effectively exploit 5GS features and QoS capabilities. The fundamentals of the logical TSN bridge are depicted in Figure 4. In the presented model, 5GS-specific procedures in CN and RAN remain hidden from the TSN network. To achieve such transparency to the TSN system and appear as any other TSN bridge, the 5GS Bridge provides TSN Ethernet ingress and egress ports on the user plane (UP) using the “Translator” functions at UE and UPF. For each pair of ingress and egress port of this logical bridge, the 5GS needs to support up to eight TSN QoS classes. Moreover, since the 5GS bridge needs to be inter-operable with the different TSN configuration models (fully centralized, hybrid, fully distributed [9]), several TSN-defined interfaces and protocols for management and control signaling need to be supported. Particularly, the TSN fully centralized configuration model requires interacting with TSN CNC for reporting of 5GS QoS capabilities and establishment of QoS flows based on received TSN stream scheduling information.

The functions required to interact with CNC are preferably located at the TSN Translator on the UPF side. Addition-

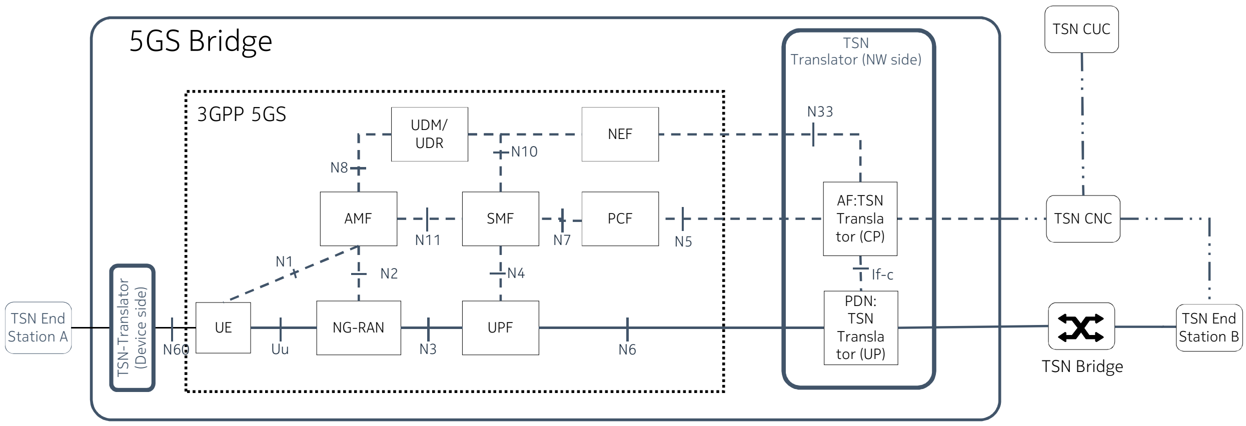

Figure 4 Integration architecture view with 5GS appearing as TSN bridge

ally, it is necessary to exchange TSN-related information between the UE-side and the UPF-side TSN Translator, e.g., using a dedicated PDU session within the 5GS. TSN related information may include, e.g., TSN configuration information, time schedules for ingress and egress ports, and time synchronization.

4 Challenges for an end-to-end integration

4.1 Radio Access Technologies

A key way to integrate 5GS and IEEE TSN is re-using the URLLC framework of 5G New Radio (5G NR). It enables minimum one-way latencies in the order of 0.5-1 ms in the RAN depending on the PHY configuration [12], and a large set of reliability enhancements to guarantee reliability levels of 99.9999 % and higher while still maintaining the radio latency below 1 ms. Assuming the jitter in the RAN is low and the core network delay is negligible, such performance is sufficient to cover many of the use cases and requirements described in Table 1. One advantage of foreseen TSN deployments, as compared to conventional URLLC use cases, is that the highly-planned and deterministic TSN traffic facilitates reservation of radio resources and latency-optimized access protocols, e.g., grant-free access including semi-persistent scheduling, without entailing large waste of resources and thus avoiding overdimensioning of the system.

Nevertheless, there are various RAN-related challenges that need to be addressed for efficient TSN support. For instance, TSN traditionally relies on time-domain access on the physical medium with transmission windows as short as 10-20 µs, whereas 5G RAN exploits frequency domain multiplexing of users on time-slots or mini-slot in the order of 100 µs or larger. This scheduling resolution mismatch calls for the use of hold-and-forward mechanisms at the edges of the 5GS (i.e., at the TSN Translators), which also need to be considered in the CP interaction with TSN CNC, i.e., as part of the packet-length independent delay.

The actual radio performance in terms of latency, reliability, and capacity depends on multiple factors. One source of variability is the time-varying UE coverage conditions as a result of mobility, and different capabilities across UEs (e.g. processing times or number of antennas). With regards to mobility, NR allows "make-before-break" handovers with zero interruption time; however, further enhancements are required to guarantee the availability of radio resources at the target cell(s) with the required latency and jitter constraints, which calls for resource reservation protocols taking into account the scheduling information provided by the TSN CNC.

4.2 Core Network and Network Management Technologies

The 5GS applies a high level of modularization in the implementation of 5GC functions. Further, control and user plane have been separated. Finally, Service-Based Interfaces (SBI) have been introduced which are used for communication between CP network functions (NFs). The CP NFs can expose their services towards other, consuming NFs. The 5GS may be composed of virtualized and non-virtualized network elements where different entities may be responsible for their management. The 3GPP network management functions are responsible for the management of the 3GPP NFs running on top of virtualized resources, as well as for non-virtualized, so-called physical network elements. In the following, we refer to the overall management and orchestration system of 5GS as "M&O" which is responsible for management of network resources and functions in virtualized and non-virtualized domains, as well as for the management of deployed network slices.

One of the fundamental challenges for the E2E integration of TSN and a 3GPP 5GS is an inherent discrepancy in their procedures. On one hand, the TSN requires an exact information about the underlying infrastructure including link speed and latencies before a TSN stream is set up. In contrast, the 5GS is able to provide such information only after the connectivity within the network has been established. Such disparity between two networks can be overcome by over-provisioning of the 5GS where the established connectivity would correspond to the most stringent requirements of the TSN counterpart. However, this is an undesirable approach due to the increased demands on the 5GS which imply resource wastage and increased costs.

The M&O system of 5GS has a complete view on the network, including the knowledge about resources, deployed services, NFs, and their performance. In addition, the M&O may contain the information about users, their requirements and characteristic. Hence, the M&O may have a fundamental role in overcoming the described mismatch, by utilizing its comprehensive knowledge about the 5GS network on the one side and the requirements of the TSN network, which usually do not change on short time scales. Therefore, the M&O can counteract the described overprovisioning by exposing estimated performance information about the 5GS to the TSN CNC and by properly configuring the 5GS.

4.3 QoS mapping between 3GPP 5GS and IEEE TSN

Regarding the 5GS Bridge integration into a TSN system using the centralized configuration model, two major challenges comprise (1) how to 派生 derive the 5GS Bridge Delay and (2) how to 执行 enforce TSN stream latency requirements within the 5GS.

Regarding (1), the definition of the TSN Bridge Delay managed object contains three indexes: ingress port, egress port and traffic (QoS) class. Similarly, the 5G QoS flows are determined by three indexes: source and destination address, and 5QI class. Since not every QoS flow in an established packet data unit (PDU) session has to carry TSN traffic, the Bridge Delay attributes of the 5GS Bridge only need to be exposed for a set of QoS flows relevant for the TSN traffic transmission. The following is assumed: (1) There is only one TSN Ethernet frame transmitted per TSN time cycle and traffic class; (2) a TSN Ethernet frame is transported in a single PDU in the 5GS Bridge; (3) consequently, there is only one PDU per data burst to be considered, i.e. the maximum data burst volume (MDBV) of the 5QI corresponds to the maximum PDU size. MDBV and maximum packet size/PDU size can be used synonymously. Following the notion of TSN Bridge Delay, the delay τ of a 5GS Bridge, corresponding to the end-to-end packet delay budget (PDB) within a 3GPP system, consists of two parts: packet size dependent and packet size independent, which can be expressed by the following linear equation:

$$\tau(P) = \frac{1}{\alpha} P + \beta, \tag{1}$$

where α represents the 3GPP guaranteed bit rate for a data flow which should be constant for each data burst, P the packet/frame size or MDBV, ($\frac{1}{\alpha}P$ is the dependent delay part) and β the independent delay part of the end-to-end 5GS Bridge delay. Note that the independent delay part reflects internal components of the 5GS Bridge such as jitter buffer, packet processing times, etc. The PDB value, being defined as an upper bound of packet delay, represents the $\tau_{\max} = \frac{1}{\alpha}P_{\max} + \beta$, i.e., PDB $= \tau_{\max}$. For each time-critical guaranteed bit rate (GBR) flow, the QoS profile incorporates the information about guaranteed UL/DL flow bit rate (GFBR), as well as maximum allowed UL/DL flow bit rate (MFBR). The GFBR and MFBR can be used as lower and upper bound value of α. Given Eq. (1), and the packet size range from 64B (minimum for TSN traffic) and $P_{\max}$, as well as the flow bit rate values, the packet size dependent part of the delay (minimum and maximum values) can be determined. Furthermore, having the packet size dependent part of the delay, the packet size independent delay β can be derived for a given QoS flow. Such 5GS Bridge delay attributes (i.e., dependentDelay and independentDelay) will be further exposed by the TSN Translator function towards the TSN CNC. In order to estimate the dependentDelay(Min/Max) values, Eq. (1) and the following input parameters shall be used:

- the information on the packet/frame size value: 64B (minimum for TSN traffic) and $P_{\max}$ which is the maximum expected packet size/MDBV for a given QoS flow indicated by e.g., the M&O entity.
- MFBR (Maximum Flow Bit Rate)
- GFBR (Guaranteed Flow Bit Rate)

Knowing the packet-size dependent delay, the packet-length independent delay β can be derived for a given QoS flow based on the linear delay expression in Eq. (1).

Figure 5 shows an exemplary relationship between performance of a 5GS Bridge implementation and reported Bridge Delay values. First, the dashed line shows the expected delay of an exemplary (but arbitrary) 5GS implementation of a TSN bridge. Apparently, this function would rather be discontinuous. The discontinuity is caused by buffer times, PDU packing/unpacking, OFDM framing, Hybrid ARQ, and other effects. The curve $\tau_{\min}(P)$ would reflect the minimum possible linear function representing the delay caused by the 5GS Bridge where $\frac{1}{\alpha}$ relates to the dependent delay and reflects the bridge's port capacity and β relates to the independent delay, which reflects static processing time of the bridge.

We now assume fixed packet sizes of P_1, P_2, or P_3, respectively. For each of these packet sizes, a minimum delay $\beta_n = \tau_{\min}(P_n)$ can be given. In 3GPP, requirements on latencies (e.g., in 5QI profiles) are expressed as absolute values and not as functions of the packet size. Therefore, a fixed value needs to be applied by the 5GS Bridge to the end-to-end connection between UE-side and UPF-side TSN Translators: the minimum delay value that can be guaranteed by the 5GS Bridge for a given maximum packet size P_1, P_2, or P_3, respectively. A lower value must not be reported because even for smaller packets than the MDBV it may happen that the full PDB is exploited by the 5GS and then the delay expected by the TSN CNC could not be met. Consequently, in a conservative approach, the following Bridge Delay parameters are reported towards the TSN CNC:

- minimum independent delay $\beta_{\min}$,
- minimum dependent delay $\frac{1}{\alpha}$,
- maximum independent delay β_n (where n depends on the assumed MDBV),
- maximum dependent delay $\frac{1}{\alpha_n}$.

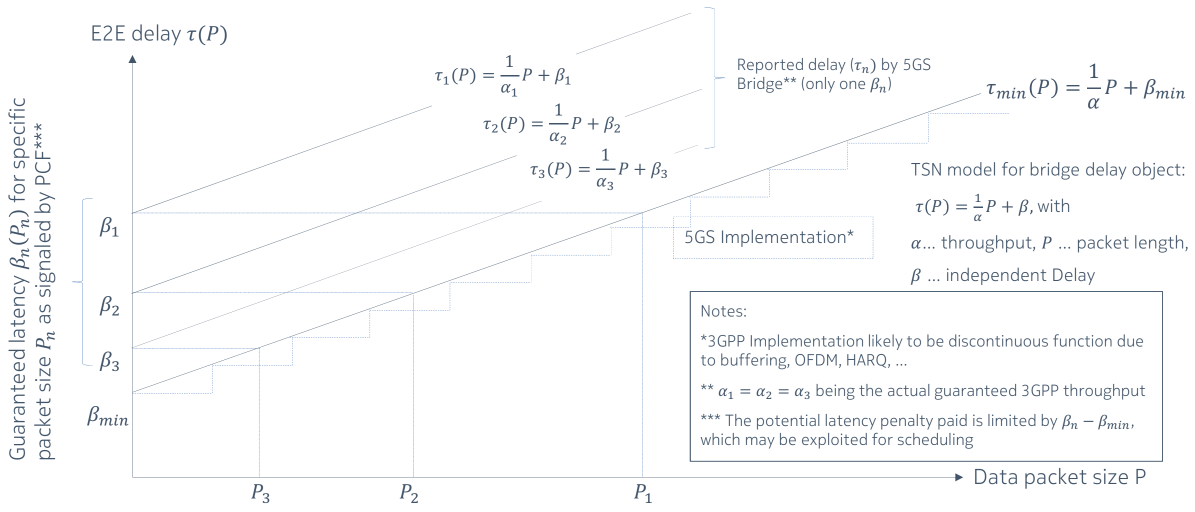

Figure 5 Exemplary relationship of 5GS implementation and reported 5GS Bridge Delays

5 Conclusions

This paper has motivated scenarios for an integrated 3GPP 5GS / IEEE TSN deployment in Industry 4.0 environments. Fundamental requirements and the 3GPP integrated system architecture have been presented, where the 5GS is modeled as a logical TSN bridge. The paper has summarized key advantages and challenges of this approach and provided a framework for QoS mapping between the 5GS and TSN systems. 3GPP will address several of these interworking aspects within the ongoing work on Rel. 16 and upcoming efforts towards Rel. 17.

Acknowledgments

The authors would like to thank their colleagues in Nokia Bell Labs. Several parts of this work are based on their efforts. Furthermore, part of this work has been performed in the framework of the project TACNET 4.0 (www.tacnet40.de), funded by the German Federal Ministry of Education and Research (BMBF). The authors would like to acknowledge the contributions of their project colleagues.

6 References

[1] H. Kagermann, W. Wahlster, and J. Helbig (Eds.), "Recommendations for implementing the strate-gic initiative Industrie 4.0: Final report of the In-dustrie 4.0 Working Group", 2013

[2] 3GPP, "TR 22.804 Study on Communication for Automation in Vertical Domains (Release 16)", Version 16.1.0, September 2018

[3] 5G Alliance for Connected Industries and Automation, "5G for Connected Industries and Automation", https://www.5g-acia.org/publications/5g-for-connected-industries-and-automation-white-paper/, April 2018

[4] "Industrial communication technology handbook," 2nd edition, Richard Zurawski, CRC Press, August 2014

[5] Avnu Alliance® White Paper, "Industrial Wireless Time-Sensitive Networking: RFC on the Path Forward", January, 2018

[6] Ahmed Nasrallah, Akhilesh Thyagaturu, Ziyad Alharbi, Cuixiang Wan,g Xing Shao, Martin Reisslein, and Hesham ElBakoury, "Ultra-Low Latency (ULL) Networks: The IEEE TSN and IETF DetNet Standards and Related 5G ULL Research", IEEE Communications Surveys and Tutorials, September 2018

[7] M. Gundall, et al. "5G as enabler for Industrie 4.0 use cases: Challenges and concepts," IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Torino, Italy, September 2018

[8] J.L.Messenger, "Time-sensitive networking: An introduction," IEEE Communications Standards Magazine, June 2018

[9] IEEE, "IEEE 802.1Q: IEEE Standard for local and metropolitan area networks – Bridges and bridged networks," May 2018

[10] 3GPP, "TS 23.501 System Architecture for the 5G System," Version 15.3.0, September 2018

[11] 3GPP, "TR 23.734 Study on 5G Enhanced sup-port of vertical and LAN services (Release 16)," Version 0.2.0, September 2018

[12] G. Pocovi, H. Shariatmadari, G. Berardinelli, K. Pedersen, J. Steiner, Z. Li, "Achieving Ultra-Reliable Low-Latency Communications: Challenges and Envisioned System Enhancements," IEEE Network, vol. 32, no. 2, March 2018

[13] 3GPP, "TR 38.825 Study on NR industrial Internet of Things (IoT) (Release 16)," Nov. 2018