http://dx.doi.org/10.25673/111641

# Performance Evaluation and Application of Real-Time Communication with 5G IIoT

Niladri Mondal[1], Dimitri Block [2], Björn Kroll[3], Florian Klingler[4]

**Abstract:** In communication systems, high data rates combined with low end-to-end latencies are prime necessities for allowing a wide variety of applications, e.g., streaming of video and data or in context of IoT Systems. In contrast, applications in Industry Automation require deterministic end-to-end latencies with guaranteed deadlines. In communication systems, data rates, reliability and the achievable end-to-end latency are often a trade-off (e.g., due to buffering of data, and overall systems-design). Further, most communication systems are optimized for high data rates only, yet, deterministic end-to-end latencies are required for most Industrial Use-Cases, which are still not considered well enough in research and standardization. In this paper we focus on low-latency communication, and, outline the importance of this research aspect. Consequently, we propose a novel Mini-Slot approach for 5G and beyond communication systems to tackle the problem of minimizing uplink- and downlink communication latencies in cellular networks under load. First evaluations of our approach in context of a feasibility study show promising results. As comparison in realistic experiments with Rel-15-based 5G Commercial off-the-shelf (COTS) hardware, a baseline scenario (unoptimized) shows a maximum latency up to 49.04 ms. In contrast to that, our novel mini-slot approach allows to lower the maximum end-to-end communication latency to 15.51 ms. This way, our mini-slot approach constitutes as enabler for low-latency communication by using Rel-15-based 5G COTS and User Equipment (UE) hardware for industrial use-cases, without the need to wait for further releases of 5G systems.

## 1 Motivation

5G New Radio (NR) is the new radio access technology developed and standardized by Third Generation Partnership Project (3GPP). It is a convergent wireless technology enabling diverse use cases in industrial applications. It can support real-time communication, which is essential for many industrial scenarios that require low latency and high reliability. One of the potential applications of 5G-based real-time communication is fieldbus communication, which are network protocols for connecting sensors, actuators, controllers and other devices in industrial automation systems. By using 5G instead of wired fieldbus, industrial applications can benefit from the wireless advantage for flexibility, mobility or motion in plants and machines in e.g. discrete manufacturing, intralogistic applications and process industry.

[1] TU Ilmenau, IoT Engineering, Ehrenbergstraße 29, 98693 Ilmenau, Germany. mondal@wnc-labs.org
[2] Weidmüller Interface Interface GmbH & Co. KG, Klingenbergstraße 26, 32758 Detmold, Deutschland. dimitri.block@weidmueller.com
[3] Fraunhofer IOSB-INA, Campusallee 1 , 32657 Lemgo, Germany. bjoern.kroll@iosb-ina.fraunhofer.de
[4] TU Ilmenau, IoT Engineering, Ehrenbergstraße 29, 98693 Ilmenau, Germany. klingler@wnc-labs.org

5G is a novel technology that defines various features to address different use cases. However, early Commercial off-the-shelf (COTS) 5G networks and devices utilize different subsets of 5G features and therefore act differently. Thus, it is essential to conduct experiments in COTS 5G networks and devices to evaluate their real-world performance and limitations.

## 2 Problem Description

A fieldbus is a communication system that connects various devices and sensors in an industrial network. There are different types of fieldbus protocols, such as PROFINET IRT, EtherCAT, Sercos III, and Ethernet Powerlink [Wo17]. These protocols have some minor differences in their data formats, addressing schemes, and error handling mechanisms, but they all require real-time performance from the physical layer of the network. This means that the data transmission and reception must meet strict timing constraints to ensure the reliability and safety of the industrial processes. Therefore, when considering the use of 5G for fieldbus applications, it is important to understand the limitations and challenges of achieving real-time communication over a wireless medium.

In order to quantify real-time communication limitations, we focus on two aspects which outline prime metrics to tackle real-time communication in an industrial context: The timing added for packet transmission and their reliability of successful transmission. The timing aspect can be measured by the One Way Latency (OWL), while the reliability aspect can be measured by the packet error rate. Therefore, by quantify and optimizing the transmission latency and the packet error rate, the real-time capability for fieldbus protocols can be validated.

A pre-configured 5G wireless communication technology is implemented in hardware setups cloning a smart warehouse scenario described in Fig. 1. The test scenario consists of a controller such as a Programmable Logic Controller (PLC) and a PROFINET RT device (CC A&B). The device is attached to one of the moving columns of the smart warehouse. Communication between the PLC and the device takes place with the help of an applied industrial 5G IIoT solution. The end-to-end latency observed in the system is to be investigated, attempted to be reduced from it's default initial value, suffering least or no loss in channel throughput.

## 3 State of the art in 5G Communication

5G was distributed under 3GPP release 15. It is a key component of the 5G standard and is designed to support a wide range of use cases and deployment scenarios. It operates on a range of frequency bands from low-band frequencies to millimeter-wave frequencies. The low-band frequencies are below 6 GHz carrier range and are also known as sub-6 GHz band or Frequency Range (FR)1. The bandwidth for FR1 ranges from 5 MHz to 100 MHz. The

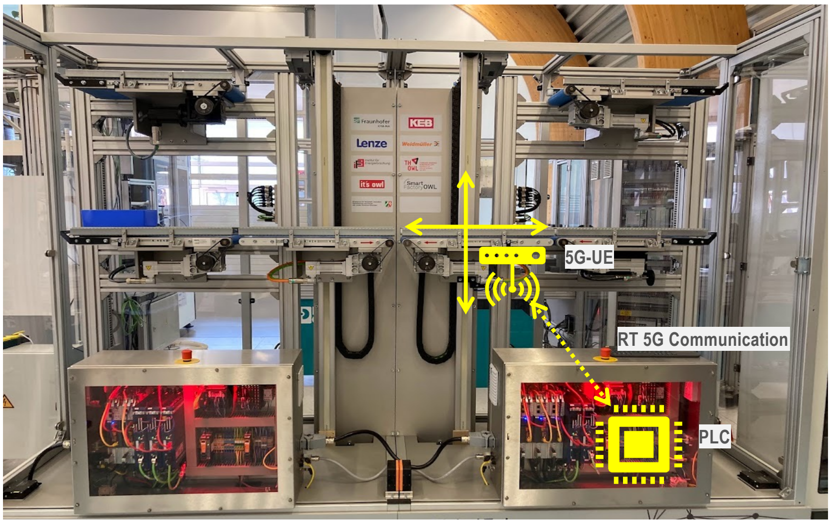


Fig. 1: Example of an IIoT Environment (Smart Warehouse) representing the problem statement of this work.

millimeter-wave band operates above 6 GHz carrier frequency range. It is also called FR2 and includes the bandwidths of 50 MHz, 100 MHz, 200 MHz, and 400 MHz.

The millimeter-waves have high frequency range and because of this, the signals can easily get degraded in quality. 5G is designed to provide higher data rates, lower latency, and more reliable connectivity than previous generations of cellular networks which enables it to deliver better performance and efficiency in wireless communication.

A 5G frame is of duration 10 ms in time which is further divided into 10 sub-frames. Each of these sub-frames contains one or more number of slots. These slots accommodate the symbols which are meant to be transmitted over the 5G wireless channel. Each of the slots in a 5G NR frame can be allotted with Downlink symbols, Uplink symbols or Flexible symbols. A symbol assigned as Flexible symbol, can be used for Downlink and Uplink traffic as per the requirement. Each slot in 5G NR can accommodate a maximum of 14 OFDM symbols.

### 3.1 Related Work

There are a large number of factory automation IIoT use cases that leverage the benefit of cellular wireless communication technologies. With respect to industrial cellular 5G technology, there have been previously researched work, where the 5G wireless network is modified to suit the latency requirements. It is based on specific use cases whether end-to-end latency is considered, or latency is closely observed only between the immediate radio communication systems, i.e within the PHY and MAC layers of the 5G base station and UE indicated by L1+L2 latency in [Wi16]. In order to reduce the observed end-to-end latency in factory automation scenarios, different communication protocols and layers have also been modified. In the work of Natale Patriciello et al. [Pa18], the IP packet size was modified. After modifying the packet size and maintaining the packet generation rate, the packet arrival rate tends to increase in the Radio Link Control buffer. Kernel RLC buffer size was modified in the work of [Ir22] to investigate the effects on end-to-end latency with an increased receiving and sending buffer sizes in the end node machines in the communication system. The solutions proposed were appropriately suited for the specific use cases. An attempt to reduce the wireless communication latency by modifying the radio frame scheduling was performed in [Lä14]. One of the main focuses of the work has been on battery powered devices where it is crucial to conserve more energy. Performance improvement at the cost of battery life can considerably affect the operational time of such devices. The work in [Lä14] was also based on consumption of low energy with the purpose of increasing the battery life and performance. The performance was compared to that of a 4G technology-based implementation. In the work of Jens Pilz et al. [Pi16], 5G network was evaluated for tactile use cases. Such application scenarios require low latency communication, generally below 1 ms with high reliability. The research was carried out using a Software Defined Radio (SDR).

## 4 Research Project 5G4Automation

The overall objective of the 5G4Automation project is to design a methodology for the development of 5G products and services in the Industry 4.0 context. This will be implemented in the form of a kit equipped with methods, guidelines and concrete implementations. This kit is intended to enable small and medium-sized enterprises in particular to develop 5G products and services on a company- and application-specific basis. In this way, a contribution is also to be made to the technical sovereignty of companies and to Germany as a technology location overall. A general overview could be found in [5G423]

## 5 Our Minislot approach in the 5G Context

In the following we outline the basic 5G System as well as our novel mini slot approach to achieve real-time communication in an industrial context.

### 5.1 5G Mini Slot Approach

5G NR allows scheduling of the 5G frame in mini-slots. A mini-slot is described as the minimum scheduling unit in 5G NR. It is an enabler of a key feature in 5G called

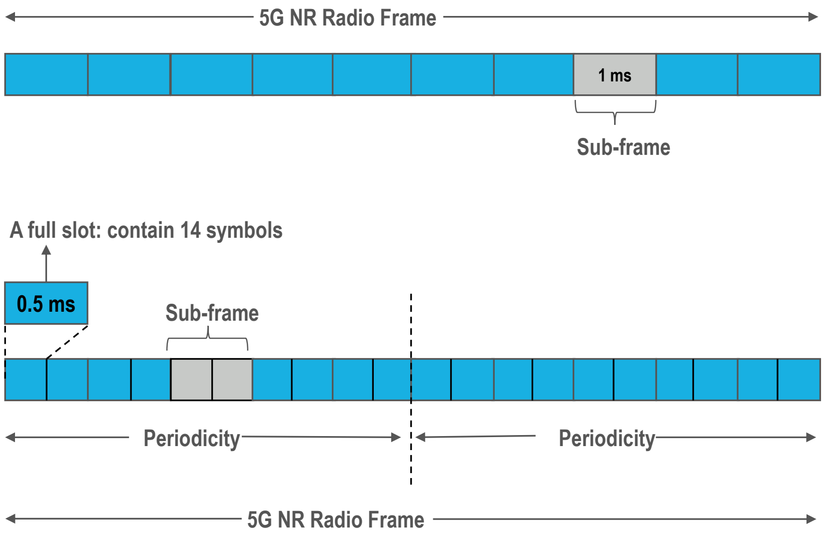


Fig. 2: The figure shows a TDD Synchronised 5G NR Frame and slot division per sub-frame at sub-carrier spacing 30 kHz used in this work.

Ultra-Reliable Low-Latency Communications (URLLC). This The mini-slot can be started at any OFDM symbol in a 5G NR slot and it can be inserted into an ongoing transmission. This enables services like URLLC along with, for example, enhanced Mobile Broadband (eMBB) traffic. Hence, it is implemented in an asynchronous mode with respect to a standard scheduled slot, which is synchronous. A mini-slot can contain either 2, 4 or 7 OFDM symbols. It can be used to transmit and receive user data in PUSCH or PDSCH channels. Hence, shorter and crucial data can be communicated using mini-slots with lower latency than that of a standard scheduled slot of 5G NR. A low latency communication provides faster and richer services because, then the 5G network can process large amounts of data in shorter time period.

This study tackles a real-time communication requiring scenario, where it is assumed that the control data for an Sensor Actuator Interface (SAI) is transmitted from a PLC over a 5G wireless channel. This is considered a Downlink (DL) network traffic. The crucial sensor data is transmitted from SAI to the PLC over the 5G wireless channel and is known as the Uplink (UL) network traffic in the problem scenario. In this regard, it is necessary that the control data from the PLC is transmitted as soon as possible to the SAI to avoid receiving any invalid or inaccurate data in the PLC. Also, the PLC is required to analyze and respond in real-time with control signals over 5G wireless network to the SAI, if there is a consequent action that needs to be performed by an actuator. In event-based systems such as PLCs and SAI, synchronous arrival of data is crucial. The operations in these systems are cyclic in nature. Therefore, if data arrives at a time that has missed the start of a cycle time duration, it is only processed at the next cycle. Consequently, a higher latency will be experienced on the actuator side. Hence, a higher end-to-end latency is realized.

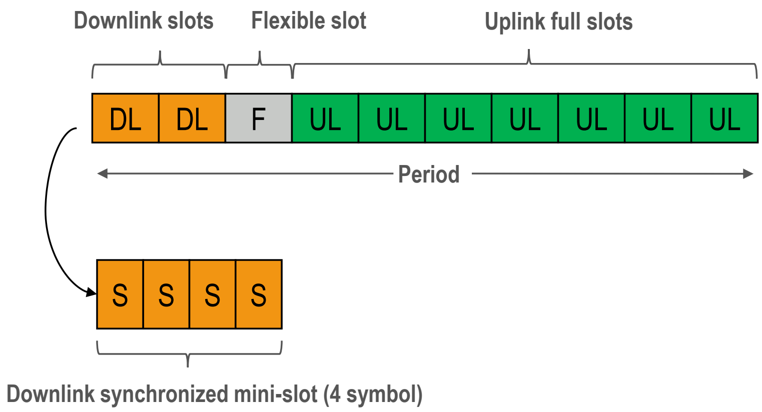


Fig. 3: TDD Synchronised 5G NR Frame containing two 4 symbol Downlink Mini-Slot

In order to solve this problem in favor of a reduced end-to-end latency and deterministic communication, mini-slots are designed in downlink direction of the network traffic in Time Division Duplex (TDD) mode. A synchronous mini-slot will introduce more deterministic traffic flow with reserved slots for transmitting control signals from the PLC to the SAI. Two downlink directed mini-slots are designed to be implemented in the 5G NR radio frame as indicated in the Fig. 3. The sensor will generate frequent data to be analyzed corresponding to the problem scenario. Hence, it is crucial that the sensor data is transmitted almost continuously to the PLC. Each of the seven uplink slots as seen in figure Fig. 3, transmit data at it's maximum capacity of 14 OFDM symbols per uplink slot. The TDD synchronized pattern uses one flexible slot. The symbols in the flexible slot can be assigned as downlink or uplink symbols based on traffic needs of 5G UE.

## 6 Real-Time Limitations of Current Industrial 5G Systems

Of the shelf 5G systems are optimized for private cellular usage patterns, like Streaming Data for example. Those system do not have a good configuration for the automation industry. As there are complete different usage patterns. Here the upload of small data frames (e.g. 128 Byte) is a good fraction of the overall traffic. Also the download speed has no value from this perspective.

In our measurements we do not focus on the achieved datarate, as for our main use case (profinet) the achieved end-to-end latency constitutes the prime metric for evaluation. To tackle this, we orchestrated sockperf and ITU-T Y.1564 for measuring the one-way latency under load.

In difference to sockperf the ITU measurements inducting additional load to the system which acts as general system load. Thus both measurements stress testing the COTS, but with different goals. Sockperf focusing on the sole system performance, ITU-T Y.1564 have the focus on shared networks.

Both measurements also outlining the limitations in current (release 15) 5G systems. As the Round Trip Time (RTT) in the worst case scenario could be more than 45 ms a lot of possible automation network types won't operate properly and need new network configurations and thus new safety evaluations.

### 6.1 Load-based Measurements with sockperf

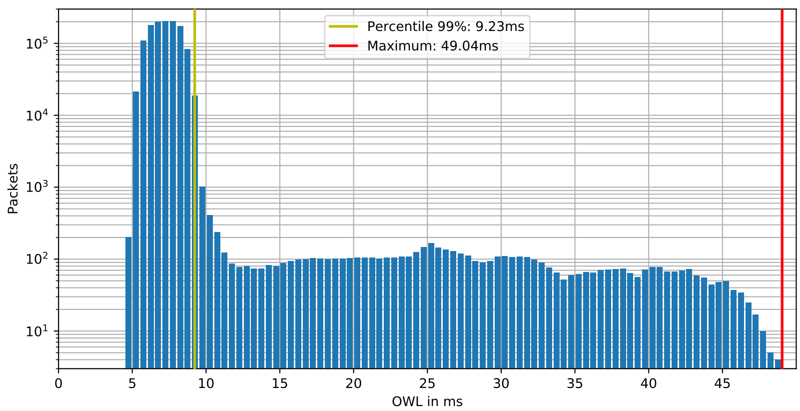


Fig. 4: Histogram of unidirectional latency measurements under load of COTS Rel-15-based 5G User Equipment (UE) and Stand Alone (SA) Non-Public Networks (NPN) (14 Byte payload with 100 µs transmission interval)

The sockperf test was performed in which for every fifth transmitted packet the server replies with a packet to emulate a utilized communication channel. This way, the measured RTT is being divided by to (as being done by sockperf) to gather the one-way latency.

The testbench consists of an Amarisoft Callbox acting as COTS and a Simcom 8200EA connected to a linux system as UE. The total runtime was 600 s which resulted in 5999964 Messages sent. The 99% percentile latency is 9.23 ms as shown in figure 4.

### 6.2 ITU-T Y.1564 Measurements

The main purpose of ITU-T Y.1564 is to provide network operators with a standardized method for measuring the end-to-end performance of Ethernet-based networks. This includes the performance of the network itself, as well as the performance of any devices or applications that are running on the network.

The test bench consists of two EXFO Ethersam ITU-T Y.1564 devices and two Mikrotik routers. The routers are connected to the 5G system as UE and the N6 interface. Both devices Internet Protocol (IP) address are static and run on the same RouterOS software (Version 7.11.2). Both devices are running as Virtual Tunnel Endpoint (VTEP) and the Ethernet ports are used for the EXFO connection as well as the connection to the 5G system.

In total three tests where performed with 128 Byte payload and additionally a 3000 Byte payload system load test (not shown). Each measurements creates 230.000 Ethernet frames to give an overview of the jitter and latency. Each Test was repeated at different signal levels to ensure a realistic view on the systems overall latency. The results are shown as an overview in Table 1. As the focus here is the mean performance each of the three measurements is averaged and summarized.

| signal strength(dBm) | OWL(ms) | jitter(ms) |
|---|---|---|
| -55 | 18.300 | 10.417 |
| -60 | 16.583 | 9.183 |
| -65 | 17.067 | 10.633 |
| -70 | 18.300 | 11.733 |
| -75 | 16.867 | 9.900 |
| -80 | 17.683 | 10.317 |
| -85 | 17.317 | 10.033 |

Tab. 1: Table of ITU-T Y.1564 measurements of COTS Rel-15-based 5G UE and SA NPN (128 Byte payload)

This measurements could taken as additional reference values for the average system performance of an of the shelf 5G system. They show, that the average OWL is 17.45 ms with a jitter of 10.32 ms.

## 7 Performance Evaluation

After making initial modifications to the 5G network as intended in section 5, we have observed an overall improvement in the performance of the 5G system in terms of OWL. The maximum OWL achieved over a test duration of 30 minutes is shown to have improved by 68.78%. For 99 percentile of the 5G network traffic, the OWL has been depicted to have reduced by 2.4%. This also points to an enhancement of reliability in 5G system used for communication scenarios requiring critical low latency.

## 8 Conclusion

We presented a novel Mini-Slot approach for scheduling downlink and uplink data transmission in a 5G and beyond cellular networking context. By taking the specific requirements of

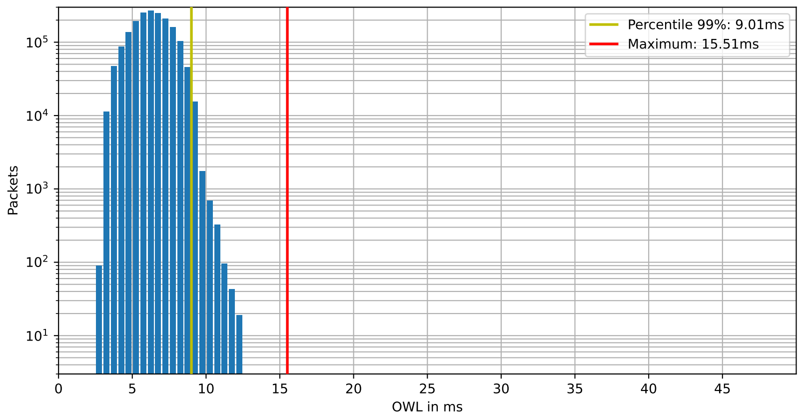


Fig. 5: Histogram of unidirectional latency measurements under load of COTS Rel-15-based 5G UE and SA NPN (14 Byte payload with 100 µs transmission interval) after modifying the 5G NR radio frame.

field bus communication for applications in the IIoT domain into account, our approach allows to lower the end-to-end communication latency on idle and under load communication links by up to 13.8% for typical industrial use-cases. Based on these results of a first feasibility study by using a Rel-15-based 5G Commercial off-the-shelf (COTS) and User Equipment (UE), our Mini-Slot approach constitutes as enabler for current and future application domains in the context of Industry Automation, where the prime requirement is low latency wireless communication.

## References

[5G423] 5G4Automation Research Project, 2023, URL: https://5g.nrw/best-practice/5g4automation/, visited on: 09/25/2023.

[Ir22] Irazabal, M.; Lopez-Aguilera, E.; Demirkol, I.; Nikaein, N.: Dynamic Buffer Sizing and Pacing as Enablers of 5G Low-Latency Services. IEEE Transactions on Mobile Computing 21/3, pp. 926–939, 2022.

[Lä14] Lähetkangas, E.; Pajukoski, K.; Vihriälä, J.; Berardinelli, G.; Lauridsen, M.; Tiirola, E.; Mogensen, P.: Achieving low latency and energy consumption by 5G TDD mode optimization. In: 2014 IEEE International Conference on Communications Workshops (ICC). Pp. 1–6, 2014.

[Pa18] Patriciello, N.; Lagen, S.; Giupponi, L.; Bojovic, B.: 5G New Radio Numerologies and their Impact on the End-To-End Latency. In: 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD). Pp. 1–6, 2018.

[Pi16] Pilz, J.; Mehlhose, M.; Wirth, T.; Wieruch, D.; Holfeld, B.; Haustein, T.: A Tactile Internet demonstration: 1ms ultra low delay for wireless communications towards 5G. In: 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Pp. 862–863, 2016.

[Wi16] Wirth, T.; Mehlhose, M.; Pilz, J.; Holfeld, B.; Wieruch, D.: 5G new radio and ultra low latency applications: A PHY implementation perspective. In: 2016 50th Asilomar Conference on Signals, Systems and Computers. Pp. 1409–1413, 2016.

[Wo17] Wollschläger, Martin and Sauter, Thilo and Jasperneite, Jürgen: The Future of Industrial Communication. In: IEEE Industrial Electronics magazine. IEEE, 2017.