Parameterizing 5G New Radio: A Comparative Measurement Study on Throughput and Delay

Simon Raffeck*, Sebastian G. Grøsvik§, Stanislav Lange§, Tobias Hossfeld*, Thomas Zinner§, Stefan Geissler*
*Chair of Communication Networks, University of Würzburg, Germany
§Norwegian University of Science and Technology
Email: {firstname.lastname}@{uni-wuerzburg.de, ntnu.no}

***Abstract*—5G New Radio (NR) is designed to support diverse services, shifting from a fixed smartphone-centric infrastructure to flexible deployment options tailored for verticals like Ultra-Reliable Low-Latency Communications (URLLC) and Machine-Type Communications (MTC). To optimize the QoS of 5G NR to the service needs, understanding the impact of the introduced configuration parameters is critical. This paper investigates the configuration space of 5G NR using open-source 5G standalone deployments based on OpenAirInterface (OAI) and srsRAN (SRS). We conduct a detailed study on the impact of Next Generation NodeB (gNB) configurations on uplink and downlink throughput and latency, we compare the 5G NR implementations of OAI and SRS, and we investigate reproducibility across different testbeds at the University of Wuerzburg and NTNU. Our datasets are made publicly available.**

***Index Terms*—5G NR, Testbed, QoS, Delay, Throughput.**

I. Introduction

The advent of 5G marks a significant leap in the evolution of wireless communication, especially due to 5G campus networks as controlled experimental platforms. As we transition into this era of advanced connectivity, it is critical to systematically analyze the implications of 5G as a whole, as well as its individual components, on fundamental performance metrics. Especially, as 5G promises significant improvements in the area of low latency applications across diverse sectors such as industrial processes, healthcare, and automation as well as significantly increased throughput over LTE deployments. The promised ultra-low latency capabilities of 5G pave the way for real-time and deterministic communication, enabling seamless interactions between machines and systems.

However, in spite of all the promises made for the next generation of mobile networks, many of the targeted performance metrics have fallen short when it comes to currently available solutions [1]. In addition, as parts of the 5G standard leave details up for individual implementations [2], it is critical to systematically investigate the current solution landscape, and to develop methodologies that enable comparing solutions against each other as well as track their progress over time. This is especially crucial as it is to be expected that different solutions will differ in their implementations to adhere to patent laws and the licenses of open-source implementations.

In this work, we explore the feasibility of achieving reproducible measurements across testbeds and gNB implementations. We conduct extensive performance measurements in open-source 5G campus deployments to quantify the impact of 5G NR on throughput and delay in both uplink and downlink directions. Our methodology and results encompass multiple system configurations and gNB implementations across two independent deployments. Additionally, we develop a measurement framework to assess the reproducibility of results across independent testbeds and gain insights into Quality of Service (QoS) metrics influenced by gNB configurations. By evaluating various configuration parameters and comparing their effects using different open source gNB implementations, we address the sensitivity of measurements to hardware and environmental differences across testbeds. We argue that the gained insights are crucial for network engineers and researchers for new 5G deployments.

To this end, we investigate the following research questions.

- RQ1 Can 5G NR measurements be effectively automated to produce reliable and reproducible results?
- RQ2 Are QoS measurements reproducible across similar but non-identical 5G campus testbeds?
- RQ3 Are QoS measurements reproducible across different gNB implementations?
- RQ4 How does gNB configuration impact the performance of 5G NR?

In answering these research questions, we make the following contributions: (i) We show that result reproducibility [3] is not given across testbeds and gNB implementations, and (ii) we identify the impact of 5G NR configuration parameters on delay and throughput. In addition, we make the obtained datasets publicly available.[1]

The remainder of this work is structured as follows. Section II covers previously conducted research in the area of 5G NR. A background and abstract model of the 5G NR and its configuration parameters are presented in Section III. Section IV details our testbed setup and measurement methodology, as well as evaluated parameter combinations. An evaluation of the obtained measurements is conducted in Section V before Section VI concludes this work.

II. Related Work

The influence of NR and the scheduling algorithms of the gNB have a significant impact on the performance of 5G deployments. To quantify this, the authors of [4] set out to provide an empirical measurement study of their Nokia 5G

[1]https://doi.org/10.5281/zenodo.13754300

978-3-903176-66-9

deployment. They provide a measurement setup and investigate the latency impact of scheduling on their gNB.

Open-source 5G deployments such as srsRAN and OAI have received attention in recent years. The authors of [5] present a stable 5G standalone testbed using srsRAN and Open5GS, including a feature comparison of the SRS and OAI gNB implementations, as well as a comparison of a number of 5G core implementations. The paper briefly presents a SRS performance trace to illustrate the data reported by SRS. However, no further performance evaluation is conducted.

The authors of [6] carry out a performance evaluation of a SRS-based 5G SA testbed. They provide insights into the elements and configurations of the testbed, and performance measurements. The theoretical performance for Round-Trip-Time (RTT) latency, Uplink (UL) and Downlink (DL) throughput is calculated and then evaluated in the testbed. The evaluations showcase a significant discrepancy between expected and actual performance for both latency and throughput. The authors identify key performance factors such as bandwidth, number of antennas, modulation and coding scheme, and frame structure. However, an exhaustive explanation for the gap between expected and measured performance remains for future investigations. The authors of [7] implement an OAI-based 5G NR simulation to compare the performance of different MAC scheduling algorithms. The results show that the choice of algorithm has an impact on system performance.

The authors of [8], [9] provide performance comparisons of 4G, 5G NSA, and 5G SA testbeds. These publications focus on the broad generational differences, by comparing the performance of the three technologies in a general manner.

In [10], [11], and [6] models for estimating the constraints of 5G NR latency performance are presented. However, the models are not unified in their estimates. For example, in the case of TDD mode with SCS of 30 kHz, in [10] the round trip latency is estimated to be 3.11 ms, in [11] 6.3 ms and in [6] either 6.5 ms or 3.5 ms depending on the SR period.

To summarize, earlier publications have tackled a variety of 5G NR performance aspects. However, there is still a discrepancy between measured and expected performance. Systematical comparative studies which investigate the impact of 5G NR configurations are largely missing. We aim to extend previous works through deeper insights into the effect of radio configuration parameters on performance. We implement a common methodology to compare implementations across multiple 5G NR options and physical testbeds.

III. Background on 5G New Radio

The frame structure in 5G NR, as shown in Figure 1, is designed for maximum flexibility, extending features from LTE radio communication [12]. 5G NR enhances radio channel configuration by replacing hard-coded parameters with dynamic ones, enabling operators to choose from various configurations. Below, we present a holistic overview of the 5G NR structure and introduce the new parameters.

In 5G NR, two duplex mechanisms are defined: Frequency Division Duplexing (FDD) and Time Division Duplexing

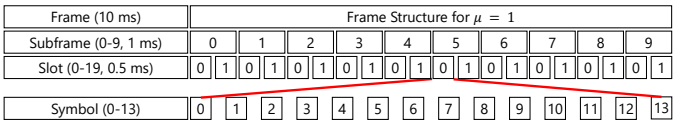

Fig. 1. Schematic representation of 5G NR frame structure ($\mu = 1$).

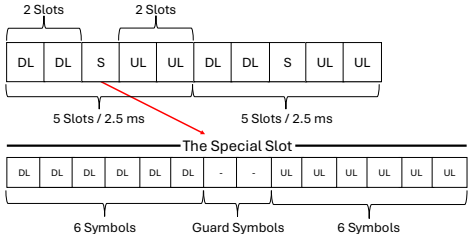

Fig. 2. Exemplary 5G NR TDD pattern with a duration of 5 slots. 2xUL, 2xDL, and one flexible slot.

(TDD). FDD uses separate frequency bands for uplink and downlink, enabling simultaneous bidirectional communication but requiring more spectral resources. TDD, however, alternates uplink and downlink transmissions within the same frequency band based on time slots, offering greater flexibility and efficient spectrum use. In this work, we focus explicitly on TDD and its configuration parameters.

A frame in 5G NR, the largest scheduling unit for TDD, has a 10 ms duration and is the fundamental time unit for TDD operation. Each frame divides into 10 subframes of 1 ms each. Subframes are further divided into slots, with the number of slots depending on subcarrier spacing, e.g., 2 slots per subframe for 30 kHz spacing. Each slot contains 14 symbols for a normal cyclic prefix and 12 for an extended cyclic prefix, as specified in 3GPP TS 38.211 [13].

The layout of individual slots in 5G NR is also configurable. The TDD pattern is highly flexible, allowing to allocate entire slots and individual symbols for uplink and downlink based on traffic demands and service requirements. This flexibility is vital for optimizing spectrum efficiency and supporting diverse use cases. The TDD pattern can be adjusted to prioritize uplink or downlink transmissions, ensuring efficient spectrum utilization and adaptability for varying user and application needs. The number of pure downlink (DL) / uplink (UL) slots can be configured, as illustrated in Figure 2. A special slot separates the DL and UL slots within the TDD pattern, and the number of UL and DL symbols within this slot can be configured. This special slot also includes guard symbols (-) to ensure a clean transition between DL and UL transmissions.

Orthogonal Frequency Division Multiple Access (OFDMA) forms the basis for resource allocation in 5G NR and is represented by a two-dimensional matrix of time and frequency resources. This matrix consists of resource blocks (RBs) in the frequency domain and slots in the time domain. A RB, the smallest frequency allocation unit, spans 12 subcarriers, while a slot is the smallest time allocation unit. Subcarriers are orthogonal to each other within a given band. The gNB dynamically assigns these RBs and slots to individual UEs based on their data rate needs, channel conditions, and

QoS requirements. This dynamic scheduling enables efficient spectrum utilization and accommodates varying network demands. The resource allocation across one frame (10 ms) in the time domain spans all available RBs and subcarriers in the frequency domain. Different UEs are assigned different RBs within the frame for data transmission. The number of assigned slots and RBs depends on the radio channel load and the data to be transmitted or received by the UE. In an isolated environment with a single UE attached to a gNB, as in the measurements for this work, the UE can be assigned all slots and RBs if sufficient data is available for transmission within a frame. Finally, the order in which slots and RBs are assigned to UEs depends on the scheduler of the gNB implementation. Current popular implementations are round-robin and proportional fair share [7]. However, there are several extensions and optimizations discussed in literature, to further increase the efficiency of the radio interface [14].

A. 5G NR Numerology

5G NR enables flexibility through dynamic numerologies which are controlled by the parameter μ. The numerology defines subcarrier spacing (SCS) and OFDM symbol duration, with support for SCS values of $2^{\mu} \cdot 15$ kHz, $\mu = \{0, .., 6\}$ [13]. The selected numerology directly impacts the slot duration and thus the OFDM symbol duration, as each slot generally contains 14 symbols. This allows fine-grained radio channel configuration for various scenarios (e.g., cell size, multiplexing) [15]. We omit further details on the impact of numerology, as only its effect on slot length is relevant for this work.

B. TDD Periodicity and Pattern

In 5G NR, the TDD pattern and its periodicity are configurable. The TDD pattern allocates time slots for uplink (UL) and downlink (DL) transmissions within a frame, allowing the network to adjust the UL-DL balance based on traffic demands. Periodicity refers to the repetition interval of a TDD pattern and must align with the selected numerology. Specifically, the periodicity must match a duration that, when repeated 1 to n times, fits precisely within a 10 ms frame. Configurable TDD patterns can vary in length and structure, supporting different traffic profiles and enabling efficient spectrum use in various deployment scenarios. Figure 2 illustrates a sample pattern with a 2.5 ms periodicity containing two UL slots and two DL slots separated by a flexible special slot. Within this special slot, operators can explicitly define the number of DL and UL symbols. Guard symbols are utilized in the remaining symbols to separate UL and DL segments of the TDD pattern, ensuring collision-free air interface operation.

IV. Methodology

A. Testbed Description

To evaluate the aspect of reproducibility, we developed a dedicated testbed using off-the-shelf components and open source software. We then deploy similar testbeds, as described in Table I, at two locations: University of Wuerzburg, Germany (UWUE), and NTNU Trondheim, Norway, enabling direct comparison of measurement outcomes. Interestingly, our analysis revealed stable performance metrics for certain configurations across both locations, while others exhibited significant variability, underscoring the influence of environmental factors on measurement outcomes. A detailed comparison of the testbeds is provided in the evaluation section.

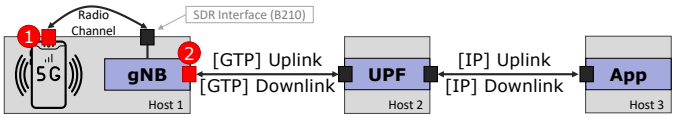

Fig. 3. Testbed setup and measurement points. (1) Radio Interface of UE (2) Backhaul Interface of gNB use the same clock of Host 1.

TABLE I
Testbed hardware setup.

	NTNU	UWUE
gNB/UE Host	Intel Core i7-6700	AMD Ryzen 7 3700X
SDR	USRP B210; UHD Version 4.0	
gNB Version	OAI: v2.1.0 (9fab212441) SRS: 24.04.0 (c33cacba7)	
Modem	Quectel RM520N-GL	
Frequency	3600-3640 MHz	

Figure 3 illustrates our testbed configuration, featuring three hosts: one each for the UE and gNB, the UPF, and an application server. To conduct all four measurement regimes — UL delay, DL delay, UL throughput, and DL throughput — it is crucial to identify suitable measurement points. Synchronized clocks are essential for accurately measuring one-way delays in both uplink and downlink directions. We achieve this by connecting a 5G modem, the Quectel RM520N-GL, to Host 1 via USB using an M.2 carrier-board, Host 1 also runs the evaluated gNB solution. This setup ensures that both the UE radio interface (1) and gNB backhaul interface (2) share the same clock, enabling precise one-way measurements. This setup facilitates the capture of ingress and egress traffic with synchronized software clocks, allowing computation of one-way delays and throughput in both directions. In preliminary experiments, we have quantified, that running both on the same host, does not introduce interference into our testbed.

For the evaluated gNB solutions, we use the two prominent open source implementations OAI [16] and SRS [17]. Measurements were consistently performed on the n79 frequency band, operating at 20 and 40 MHz bandwidth. A dedicated private network connects Hosts 1, 2, and 3, ensuring a 1 Gbps transmission path between hosts without interference from cross-traffic that could affect measurements.

B. gNB Configurations

After introducing the configuration parameters in the background and detailing the testbed setup, we now discuss specific parameter combinations included in the study. We first identified settings to control each parameter for both SRS and OAI, noting significant differences in parameter specification methods. For instance, SRS defines bandwidth in MHz, while OAI uses the number of resource blocks. Similarly, OAI configures periodicity in milliseconds, whereas SRS uses slot numbers. Table II summarizes the evaluated parameter combinations,

TABLE II
EVALUATED PARAMETER CONFIGURATIONS ACROSS BOTH OAI AND SRS FOR BOTH 20 AND 40 MHZ.

TDD Period	DL:UL Ratio	DL Slots	DL Symbols	UL Slots	UL Symbols	μ
20 slots	1:1	10	5	9	7	1
20 slots	2:1	13	5	6	7	1
20 slots[a]	4:1	15	12	4	0	1
10 slots[a]	1:1	5	5	4	7	1
10 slots	2:1	6	8	3	4	1
10 slots	4:1	7	12	2	0	1
5 slots	1:1	2	6	2	6	1
5 slots	2:1	3	5	1	7	1
5 slots	4:1	3	12	1	0	1

[a]Only evaluated for SRS as OAI crashes

including TDD period, UL-to-DL slot ratio, and distribution of UL and DL symbols in the special slot. The numerology remained fixed at $\mu = 1$ to maintain parameter consistency.

1) TDD Periodicity: When it comes to the configuration of the TDD periodicity, the available values are fully determined by the chosen numerology of $\mu = 1$, as the numerology dictates the slot duration of $0.5\,\mathrm{ms}$. Based on this, valid values for the periodicity need to be chosen so that by repeating the period between one and n times, a frame of 20 slots and $10\,\mathrm{ms}$ duration is filled. Hence, valid values are a period of 20 slots (1 repetition à $10\,\mathrm{ms}$), 10 slots (2 repetitions à $5\,\mathrm{ms}$), and 5 slots (4 repetitions à $2.5\,\mathrm{ms}$), as shown in Table II. The expected impact of this configuration is that shorter periods (fewer slots) reduce latency, as the gNB has the chance to reassemble and deliver packets after each period, if a full GTP-encapsulated IP packet has been received. A shorter period means that this process can happen more often.

2) DL-UL Ratio and Special Slot Configuration: After deciding on a periodicity, the slots can be configured as either uplink or downlink, separated by the special slot. We compare three layouts corresponding to DL slot ratios of 1:1, 2:1, and 4:1. To avoid inflating the parameter space, we do not vary the special slot layout but use it to get configurations as close to the specified ratio as possible. Thus, some configurations have only DL symbols in the special slot, while others have a more even distribution. The expected impact is that the total number of DL and UL symbols correlates strongly with achieved throughput, as more resources in both time and frequency domains are dedicated to transmitting data in either direction.

C. Measurement Procedure

In total, we conduct four measurement regimes to provide a holistic performance overview: (1) Uplink One-Way-Delay (OWD) (2) Downlink OWD (3) Uplink throughput (4) Downlink throughput. We use Ansible scripts to fully automate the measurements and eliminate the impact of timing and ensure consistency between repetitions. The script randomizes the order of the measurement repetitions to prevent result contamination and counter environmental impacts as best as possible. During each measurement run, we capture packet traces at the UE interface towards the radio channel and on the gNB backhaul interface as shown in Figure 3.

For OWD measurements, we generate a constant bitrate stream with deterministic inter-arrival times of $1.7\,\mathrm{ms}$, chosen to avoid synchronization with the gNB's packet reassembly process occurring every TDD period ($5\,\mathrm{ms}$, $10\,\mathrm{ms}$, or $20\,\mathrm{ms}$). To mitigate pattern formation from deterministic inter-arrival times, we also perform measurements using negative exponential distributions, though these are generally omitted in the evaluation due to similar results. Each parameter combination in Table II undergoes 5 repetitions, transmitting 10,000 packets in both upstream and downstream directions. For the throughput measurements, we again use Ansible playbooks to ensure consistency and use iperf3 in UDP mode to conduct 5 repetitions for each parameter combination.

V. EVALUATION

Using the testbed and methodology described, we extract two datasets—one for OWD and one for throughput—to analyze the impact of configuration parameters on gNB performance. For uplink delay, timestamps are collected at the UE and gNB egress, and for downlink OWD at the ingress of both nodes. These timestamps are used to calculate packet inter-arrival times to validate traffic generation and measure delays. For throughput, *iperf3* logs throughput and jitter at $1\,\mathrm{s}$ intervals, and values from both server and client sides determine uplink and downlink speeds. Pre-processing filters out repetitions with loss, removing 41 of 500 runs.

We evaluate results using a top-down approach, comparing NTNU and UWUE testbeds, with results primarily from NTNU unless stated otherwise.

A. Testbed Comparison

To analyze the measurements across the two testbeds, we compare delay and throughput values. Table III presents the mean uplink and downlink throughput for OAI and SRS. For UWUE, two setups with different elevations between the gNB and UE were investigated, denoted as (1) for elevated gNB and (2) for identical elevations. Each cell shows the mean throughput and 95% confidence interval, with color indicating overlapping intervals (green) or statistically significant differences (red). Downlink and uplink behaviors differ across gNB implementations. SRS shows consistent downlink results for both setups at UWUE, but uplink data reveals notable differences. The impact of DL:UL ratio configuration is evident, with throughput scaling according to slot allocation for SRS. In contrast, OAI exhibits consistent uplink but significant downlink differences. Similarly, the trend due to the number of available downlink slots does not hold true for OAI, as a ratio of 4:1 produces lower throughput than a ratio of 2:1.

For OWD comparison, the mean difference of the mean OWD is used as metric. OAI showed significant differences in 8 of 32 DL and 4 of 32 UL scenarios, while SRS showed significant differences in 8 of 36 DL and 4 of 36 UL scenarios. The throughput and delay results reveal significant performance variations between gNB implementations across testbeds in some cases, while others yield consistent outcomes. Despite using identical gNB configurations and exclusive

TABLE III
COMPARISON OF MEAN THROUGHPUT AND 95% CONFIDENCE INTERVAL FOR 20 MHZ IN BOTH TESTBEDS FOR OAI AND SRS. TWO PHYSICAL LAYOUTS FOR UWUE TESTBED.

gNB	Ratio	DL (Mbps)			UL (Mbps)		
		NTNU	UWUE (1), (2)		NTNU	UWUE (1), (2)	
SRS	1:1	42.81±1.03	42.33±1.34	42.87±1.04	8.96±3.90	3.30±0.70	18.05±2.23
	2:1	55.64±0.43	55.77±0.52	55.54±0.40	5.04±2.72	3.25±0.42	17.21±1.61
	4:1	59.86±3.50	60.62±3.85	60.02±3.52	5.94±3.14	2.11±0.42	13.98±0.10
OAI	1:1	42.20±0.84	39.12±1.97	38.37±1.83	11.31±4.42	11.66±4.16	9.51±3.37
	2:1	55.10±4.21	49.66±3.68	49.51±4.04	8.47±1.29	7.80±1.11	7.17±1.11
	4:1	39.25±1.13	36.42±1.24	35.87±0.80	8.25±0.07	7.90±0.08	7.13±0.10

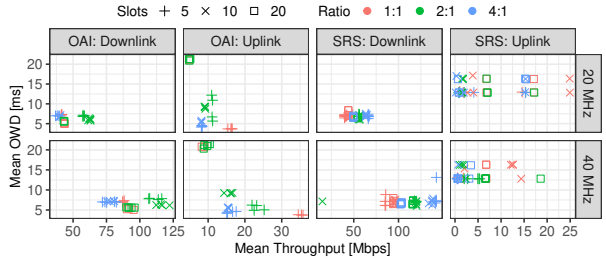

Fig. 4. Overview of delay and throughput values for different configurations.

frequency bands in both testbeds, the gNBs behave differently, with SRS even producing distinct results after physically moving the gNB. Especially for the latter, the large increase by almost 600% between scenario (1) and (2) highlights the high sensitivity against external factors, such as positioning of gNB in relation to the UE. Finally, there are certain impact factors we did not account for during our measurements, such as environmental factors, interference with other radio signals, etc. In order to further study these impact factors in the context of open-source 5G deployments, additional measurements and experiments are required, which remain for future work.
Based on the observations, we can answer RQ2: consistent results were not reproducible across the two testbeds. While similar trends appeared in both delay and throughput measurements, statistically significant differences were found in 24 out of 136 evaluated parameter combinations.

B. Performance Overview

Figure 4 presents delay and throughput for both OAI and SRS in various configurations. Each combination of color and shape is used to represent a slot and ratio configuration, as discussed in Table II. On the x-axis, the mean throughput for the respective direction is shown. On the y-axis, the mean One-Way-Delay is depicted. Note that individual facets are scaled differently along the x-axis to better visualize the differences. For downlink throughput, both implementations exhibit clusters for each of the configurations, showing an impact of slot and symbol ratio. The throughput values for 20 MHz have already been discussed in Table III. For 40 MHz, we see a steady improvement for SRS and OAI with a comparable influence of the slot and ratio parameters. SRS increases its downlink throughput to between 80 Mbps - 140 Mbps and OAI to 75 Mbps - 125 Mbps, depending on the TDD pattern, and hence the number of available downlink symbols.

The uplink throughput sees no significant improvement for SRS. For OAI the uplink throughput greatly increases, reaching a maximum of around 35 Mbps. Furthermore, the TDD pattern has a significant impact on the uplink performance, with 5 slots and a 1:1 ratio showing the best results and 20 slots with a 1:1 ratio the worst performance, even more so in latency. A trend can be seen, that the uplink throughput increases with fewer slots and a ratio closer to 1:1.

For the OWD, the data shows less influence of different configurations than for throughput. The observed clusters do not hold true for uplink and variance is much higher. The uplink communication exhibits far more outliers for both metrics. Especially for OAI, the 20 slot configurations appear less stable, reaching a maximum of 21.6 ms OWD. The SRS uplink exhibits two clusters for OWD at roughly 12.8 ms and 17 ms. Throughput scales well with available downlink slots, symbols, and bandwidth for both SRS and OAI. Downlink delay remains stable across all configurations, indicating that periodicity, TDD pattern, or bandwidth are not limiting factors. Due to the transmitted data being small (8B payload), delay is only affected if a packet can't be sent in one TDD period. Since our test device consistently uses all RBs, the measurement packet can always be transmitted in the next assigned downlink slots even at 20 MHz. Uplink delays differ between implementations: OAI shows a wide range, while SRS exhibits bi-modal behavior, mostly independent of configuration. OAI's outliers with a 4:1 slot ratio are explained by packets taking more than one pattern due to insufficient uplink symbols. The bi-modality of SRS uplink delay is explored in Figure 7.
The testbed and methodology effectively address RQ1, enabling reliable automation of 5G NR measurements with 500 tests conducted across two distinct testbeds. Additionally, the data reveals systematic differences between the two gNB implementations in throughput and delay, offering partial insights into RQ3. The data also allows us to answer RQ4, showing a systematic impact of configuration parameters on observed QoS values.

C. Throughput

Figure 4 has already shown the impact of the different slot and ratio configuration on the achieved throughput. To further quantify this impact, we compare the observed measurement values for both evaluated gNB implementations.

In Figure 5, the x-axis shows each configuration, with the colors depicting the different traffic directions. The y-axis is used to present which of the two implementations outperforms the other in regard to mean throughput, with the difference given in Mbps. Thereby, positive values indicate a higher throughput for OAI, while negative values favor SRS. The whiskers represent the 95% confidence interval. It is important to note that these results hold true in our system, but may change in a different environment or setup and should therefore only be taken as tendencies. Generally, OAI appears to outperform SRS when it comes to uplink throughput. This is in line with the observations presented in Section V-A, indicating that OAI is more resilient to external factors in uplink direction than SRS. The differences are minor except for the 5 slot 1:1 ratio scenario with 40 MHz, where the difference shows more dominantly. In two configurations only,

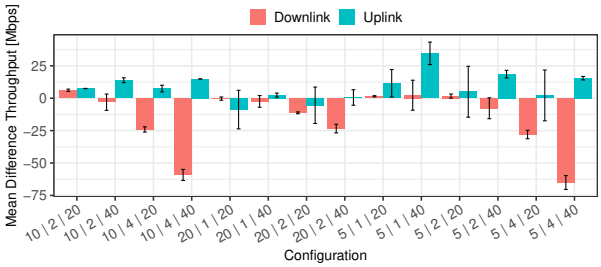

Fig. 5. Mean difference of throughput across five repetitions and various configurations. Positive: OAI higher throughput than SRS.

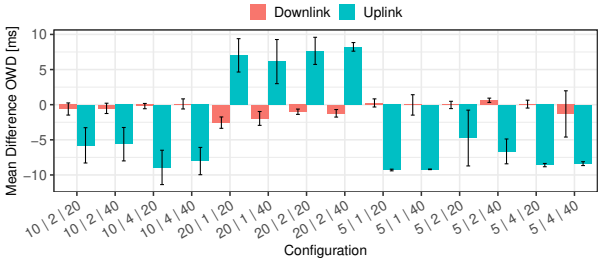

Fig. 6. Mean difference of one-way-delay across five repetitions for each configuration. Positive: OAI higher delay than SRS.

SRS comes out slightly on top for uplink traffic. For downlink throughput, SRS performs consistently better overall, only being slightly behind in 4 configurations. However, a close investigation of the confidence intervals, shows that for many of the configurations either of the implementations comes out on top. This can be explained by the influence of the radio interface itself. Different outside influences, like interference in single runs, has a huge impact on the mean throughput and can not easily be quantified for this investigation.

In summary, the throughput measurements allow us to further solidify the answer to RQ3, showing that 17 of the 28 comparisons exhibit statistically significant differences between gNB implementations, even when investigating a single testbed, under identical circumstances for both implementations.

D. Delay

The same comparison method is used in Figure 6, which displays the OWD performance difference on the y-axis. A larger bar indicates higher delays for a specific implementation. The figure shows that OAI exhibits worse OWD than SRS in four uplink configurations and slightly worse in two downlink configurations. In all other configurations, SRS has higher latency, particularly in uplink. The large confidence intervals highlight the influence of external factors on gNB performance. The repetitions within each configuration are analyzed in detail. For OAI, the focus is on delay values with 5 slots and a 2:1 ratio at 20 MHz. Downlink shows a consistent 99% quantile of around 11 ms latency, with median OWDs around 7 ms and 25% quantiles at 5 ms, exhibiting stable distributions across runs. However, for uplink, while 60% of OWDs are at most 5 ms, distributions vary significantly beyond that, with a 99% quantile for maximum OWD ranging from 32 ms to 57 ms, indicating less stability. In Figure 7 the

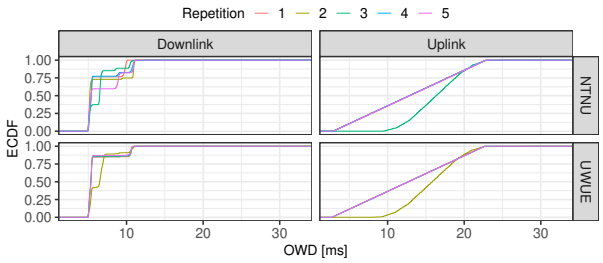

Fig. 7. Delay distribution for SRS across multiple repetitions with 20 slots, ratio 2:1, and 20 MHz.

CDFs are shown for SRS for the configuration with 20 slots, a 2:1 ratio, and 20 MHz bandwidth. The downlink OWD for SRS appears less consistent than OAI. Even though each of the runs converges towards a 99% quantile of around 11 ms, the various other quantiles vary significantly more than was observed for OAI. However, a systematic behavior can be identified. While every run appears to behave differently, each run depicts the same discretization that presents as different steps in the CDF, indicating that SRS transmits packets in the downlink at discrete points in time, or in fixed intervals. However, additional measurements and a more thorough code review is required to validate or refute this assumption.

For the uplink traffic, most of the runs exhibit a uniform distribution, ranging from 3 ms to 22 ms. Some runs, start the range with a slight offset from 8 ms, but exhibit a 99% quantile of 22 ms. To verify that this behavior can be considered systematic, the second facet shows the CDF for the same configuration captured in the setup at UWUE. The same discretization pattern in downlink and the offset in the uplink curve can be observed, suggesting a systematic behavior rather than outliers. The timeseries for one of these runs confirms these assumptions. The same saw-tooth pattern as in every other run can be observed, however in this case, the baseline is offset to a minimum of 8 ms rather than 3 ms.

These observations tie in neatly with the clusters for the mean values seen in Figure 4, which were located at around 12.8 ms and 17 ms. The offset curve, thus, correlating to the 17 ms cluster and the uniform distribution to the 12 ms.

To better understand this systematic behavior observed in both testbeds, a single run from the displayed CDFs is analyzed further as a time series. Figure 8 shows a subset of the sent packets in one exemplary measurement run using SRS. On the x-axis, a subset of the transmitted packets identified by their sequence number is depicted. On the y-axis, the measured OWD for the corresponding packet is shown. The downlink traffic mostly assumes two distinct values for the OWD. The first being around 5 ms and the other being at around 9 ms. This correlates with the CDF for run 5 (pink), shown in Figure 7. There are a few outliers with a slightly higher OWD, but these two values remain stable throughout, with each being assumed by a number of subsequent packets.

The uplink traffic exhibits another interesting behavior. The values oscillate between 3 ms and 23 ms, exhibiting a clear sawtooth pattern. Further analysis shows that each sawtooth

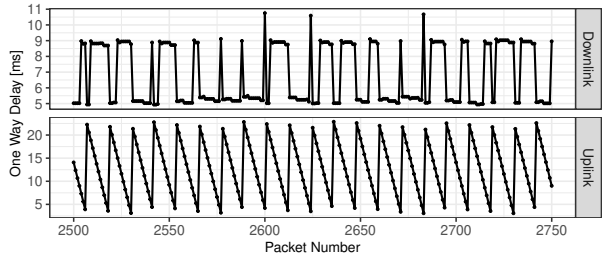

Fig. 8. Single run timeseries for SRS with 20 slots, ratio 2:1, and 20 MHz.

consists of a mean of 11 packets, with a steadily declining OWD by $1.5\,\mathrm{ms}$ on average. This sawtooth pattern is exhibited by most of the SRS configurations in some variation. However, the observed metrics for the pattern remain largely consistent for all evaluated configurations. This is once again in line with the observations made in Figure 4 that show consistent delays across all configurations for SRS. While not as pronounced as with SRS, OAI also systematically exhibits characteristic sawtooth patterns for both up- and downlink directions.

In summary, aligning with throughput observations, the answer to RQ3 is reinforced, as data shows statistically significant differences between gNB implementations in 18 of 28 configurations. Notably, the uplink exhibits more pronounced differences, while downlink delay, though statistically significant, shows smaller variations. This suggests possible differences in resource scheduling between the implementations, meriting further investigation.

VI. Conclusion

As the global 5G rollout advances and 6G testbeds and research emerge, understanding the impact of various factors on expected Quality of Service (QoS), particularly regarding the reliability and reproducibility of measurements, becomes increasingly critical. To ensure high-quality research, we conducted a detailed parameter study on the impact of various 5G New Radio (NR) configuration parameters and evaluated the reproducibility of measurements across different testbeds and open-source Next Generation NodeB (gNB) implementations. With over 500 measurement runs, we have evaluated the effect that TDD period and pattern, slot configuration as well as bandwidth have on both uplink and downlink throughput and delay. We presented our methodology to conduct One-Way-Delay (OWD) measurements for both uplink and downlink between the User Equipment (UE) egress and gNB egress. This addresses RQ1, showing that reliable automation of 5G NR measurements across setups is achievable. By applying our methodology to two independent testbeds, we addressed RQ2 and RQ3, demonstrating that reproducibility is not guaranteed across testbeds or gNB implementations, as identical configurations can lead to statistically significant differences in results. This observation holds true for both throughput and delay measurements, and marks our first major contribution. The obtained data was used to answer RQ4, by clearly showing the impact of the various configuration parameters on the measured delay and throughput values, showcasing the strong configurability of 5G NR for different use cases and providing our second main contribution. Finally, the differences across testbeds and implementations highlight the need for more in-depth performance evaluations. We identified that reproducible research requires detailed information on all components, configurations and environmental influences.

Acknowledgments

This work was partially funded by the Research Council of Norway through the SFI Norwegian Centre for Cybersecurity in Critical Sectors (NORCICS) project no. 310105 as well as ORIGAMI project from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101139270.

References

[1] M. Ghoshal, Z. J. Kong, Q. Xu *et al.*, "An in-depth study of uplink performance of 5G mmWave networks," in *Proceedings of the ACM SIGCOMM Workshop on 5G and Beyond Network Measurements, Modeling, and Use Cases*, 2022.

[2] B. Khan, A. Mihovska, R. Prasad *et al.*, "A Study on Cross-Carrier Scheduler for Carrier Aggregation in Beyond 5G Networks," in *URSI Atlantic and Asia Pacific Radio Science Meeting*. IEEE, 2022.

[3] H. Plesser, "Reproducibility vs. Replicability: A Brief History of a Confused Terminology. Front Neuroinform. 2017; 11: 76," 2017.

[4] J. Rischke, C. Vielhaus, P. Sossalla *et al.*, "Empirical Study of 5G Downlink & Uplink Scheduling and its Effects on Latency," in *IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2022.

[5] L. Mamushiane, A. Lysko, H. Kobo *et al.*, "Deploying a Stable 5G SA Testbed Using srsRAN and Open5GS: UE Integration and Troubleshooting Towards Network Slicing," in *Int. Conf. on Artificial Intelligence, Big Data, Computing and Data Communication Systems*, 2023.

[6] J. E. Håkegård, H. Lundkvist, A. Rauniyar *et al.*, "Performance Evaluation of an Open Source Implementation of a 5G Standalone Platform," *IEEE Access*, 2024.

[7] R.-M. Ursu, A. Papa, and W. Kellerer, "Experimental Evaluation of Downlink Scheduling Algorithms using OpenAirInterface," in *IEEE Wireless Communications and Networking Conf. (WCNC)*, 2022.

[8] M. Chepkoech, N. Mombeshora, B. Malila *et al.*, "Evaluation of Open-Source Mobile Network Software Stacks: A Guide to Low-cost Deployment of 5G Testbeds," in *18th Wireless On-Demand Network Systems and Services Conf. (WONS)*, 2023.

[9] R. Mihai, R. Craciunescu, A. Martian *et al.*, "Open-Source Enabled Beyond 5G Private Mobile Networks: From Concept to Prototype," in *Int. Symp. on Wireless Personal Multimedia Communications*, 2022.

[10] Y. Zhao and W. Xie, "Physical Layer Round Trip Latency Analysis and Estimation for 5G NR," in *International Wireless Communications and Mobile Computing (IWCMC)*, 2023.

[11] J. Sachs, G. Wikstrom, T. Dudda *et al.*, "5G Radio Network Design for Ultra-Reliable Low-Latency Communication," *IEEE Network*, 2018.

[12] D. Hui, S. Sandberg, Y. Blankenship *et al.*, "Channel coding in 5G new radio: A tutorial overview and performance comparison with 4G LTE," *IEEE Vehicular Technology Magazine*, 2018.

[13] 3rd Generation Partnership Project (3GPP), "3GPP TS 38.211 V15.4.0 (2019-06): NR; Physical channels and modulation (Release 15)," 3GPP, Tech. Rep., 2019.

[14] A. M. Nor, O. Fratu, S. Halunga *et al.*, "Demand based Proportional Fairness Scheduling for 5G eMBB Services," in *International Black Sea Conf. on Communications and Networking*, 2022.

[15] A. A. Zaidi, R. Baldemair, V. Molés-Cases *et al.*, "OFDM numerology design for 5G new radio to support IoT, eMBB, and MBSFN," *IEEE Communications Standards Magazine*, 2018.

[16] N. Nikaein, M. K. Marina, S. Manickam *et al.*, "OpenAirInterface: A Flexible Platform for 5G Research," *SIGCOMM Comput. Commun. Rev.*, 2014.

[17] I. Gomez-Miguelez, A. Garcia-Saavedra, P. D. Sutton *et al.*, "srsLTE: an open-source platform for LTE evolution and experimentation," in *Int. Workshop on Wireless Network Testbeds, Experimental Evaluation, and Characterization*, 2016.