M. Einhaus, M. B. Charaf, I. Kim and P. Arnold, "Bandwidth Part Adaptation and Processing Time Evaluation with OpenAirInterface," 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 2018, pp. 1-5.

doi: 10.1109/VTCFall.2018.8690902

Bandwidth Part Adaptation and Processing Time Evaluation with OpenAirInterface

Michael Einhaus[1], Mohamad Buchr Charaf[1], Igor Kim[1], and Paul Arnold[2]
[1]Hochschule für Telekommunikation Leipzig (HfTL)
einhaus@hft-leipzig.de, s169103@hft-leipzig.de, gs2016007@hft-leipzig.de,
[2]Deutsche Telekom AG
paul.arnold@telekom.de

Abstract— **Network function virtualization (NFV) and cloud based radio access network (C-RAN) architecture concepts constitute an essential foundation of the 5G mobile radio network design. Especially the envisioned splitting of RAN functionality into distributed units (DUs) at base station sites and centralized units (CUs) running in a centralized cloud offer significant advantages regarding energy efficiency, computational elasticity and cost reduction for mobile network operators (MNOs). This paper presents an exemplary implementation and initial evaluation results of a corresponding interface between DU and CU for an emulation of 5G New Radio (NR) based bandwidth part adaptation and related physical layer processing time monitoring in LTE eNBs. Such an interface will facilitate computational elasticity by means of processing time aware transmission parameter adaptation. The implementation is an extension of the FlexRAN framework for OpenAirInterface.**

Keywords—5G, radio access network (RAN), bandwidth parts (BWPs), processing time evaluation, software defined networking (SDN), network function virtualization (NFV), functional splitting, computational elasticity

I. INTRODUCTION

The overall objectives of 5G mobile radio networks encompass stringent performance requirements in terms of throughput, latency, spectral efficiency [1]. The additional aspiration for energy efficiency and high flexibility in the presence of significantly increased degrees of diversity in use cases and services, ranging from latency critical applications such as autonomous driving and smart factories to enhanced mobile broadband services with high throughput requirements, puts further pressure on the system design and operation.

The utilization of software defined networking (SDN) and network function virtualization (NFV) principles are therefore essential features of the next generation radio access network (NG-RAN) design for leveraging envisioned network slicing concepts [2]. These concepts are taken into account from the very beginning of development and standardization activities in that area [3].

The purpose of having software defined RAN (SD-RAN) architectures is providing a simplified and flexible way of base station coordination for improving the spectral efficiency by means of coordinated resource allocation and scaling of system capacity. Especially the employment of advanced interference coordination strategies will become an extremely important asset in future 5G deployments due to inherent network densification and massive MIMO concepts with comprehensive beam management necessities. Software based concepts furthermore facilitates enhancements and extensions of RAN functions by means of software updates in a flexible and efficient fashion.

Combining virtualization of RAN based functionality with cloud technology concepts is expected to provide further benefits by increasing energy efficiency and reducing total cost of ownership (TCO) in terms of capital expenditure (CAPEX) and operational expenditure (OPEX) [4]. Such kind of cloud based architecture will enable for example advanced concepts such as RAN as a Service (RaaS) that leverage cloud technologies to implement functional splits in 5G mobile radio networks [5].

The first specifications for the 5G New Radio (NR) air interface and the corresponding RAN architecture have just been published in December 2017, incorporating concepts motivated by the virtualization and cloudification paradigm. Special emphasis is thereby put on the functional split between distributed and centralized units (CU and DU, respectively) within logical base station entities [6]. As investigated for example in [7], the CU/DU split has to be conducted under accurate consideration of processing time requirements of individual functions within a base station.

In this paper, we present an exemplary implementation and extension of an interface between CU and DU based on OpenAirInterface (OAI) [8] and the corresponding FlexRAN implementation [9] which provides an extensible SDN framework for SD-RAN concepts.

The interface between CU and DU is used for the configuration and adaptation of air interface transmission parameters of the LTE eNB implementation running in a DU, and for the reporting of the processing time of individual physical layer (PHY) functions within the DU to the CU. This combination of processing time reporting and transmission parameter adaptation facilitates the application of computation elasticity strategies that take into account the utilization of computational resources (processor utilization, memory utilization, etc.) during transmission parameter optimization.

The interface extension presented in this paper is used for an exemplary initial performance study addressing the relation between processing time requirements in the downlink transmission chain of a software based eNB implementation and the achievable downlink throughput under consideration of bandwidth part configurations. The evaluation of specific algorithms for dynamic adaptation of transmission parameters depending on the processing time evaluation is not focus of this paper.

Section II provides an overview of OAI and FlexRAN in the context of concept development for RAN virtualization in cloud environments. The test environment that has been used for the performance study is described in this section as well. The developed interface extension is presented in Section III. The results of the conducted performance evaluation are discussed in Section IV and concluding remarks are provided in Section V.

II. OPENAIRINTEFACE AND FLEXRAN

A. OpenAirInterface

The OAI Software Alliance [8] provides a comprehensive open-source development environment for software defined radio (SDR) incorporating concepts such as SDN and NFV. The intention of the OAI project is the development of a real-time 5G protocol stack for running on commercial off-the-shelf (COTS) hardware. In this paper, we work with the LTE implementation of OAI.

The general setup of an OAI based system consists of an evolved packet core (EPC) and a standard conform eNB implementation.

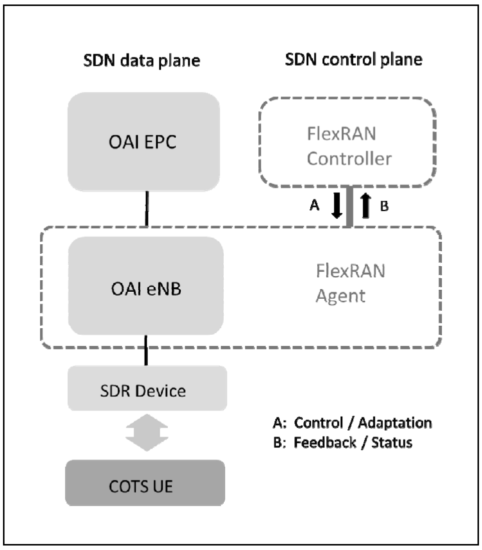

Fig. 1. FlexRAN based test setup

B. FlexRAN Concept

FlexRAN [9] is an open-source project that provides a flexible and extensible SD-RAN platform. It has been developed with the intention of providing researchers and developers a reference for evaluation of SD-RAN concepts. For that purpose, the framework basically extends the OAI implementation with SDN functionality.

Following the SDN paradigm of separating control and data plane [10], the FlexRAN implementation introduces FlexRAN Agents (FR-A) and a FlexRAN Controller (FR-C), where the latter represents the SDN controller

The communication between FR-C and FR-A forms the control plane while the LTE data traffic flow within the OAI components (EPC, eNB) represents the data plane within the SDN context. The asynchronous communication between FR-C and FR-A is facilitated by a proprietary FlexRAN protocol which uses Google Protocol Buffer [11] for message implementation and parsing.

In the context of the CU/DU split, that is currently discussed at 3GPP as one of the essential enablers for meeting the stringent RAN requirements, CU and DU would correspond to FR-C and FR-A, respectively.

The FlexRAN implementation version[1] that has been used within the scope of work presented in this paper is tightly coupled to the OAI eNB implementation. One of the key concepts is here that RAN related statistics are periodically reported from the FR-A to the FR-C which are then stored in a RAN Information Base (RIB) within the FR-C [9].

C. Test Setup

The test setup that has been used for the performance study is shown in Fig.1. It consists of three computers with four cores (4 GHz) for EPC, FR-C (representing the CU), and FR-A (representing the DU), respectively. All computers run with Ubuntu 16.04 and Linux kernel version 4.10.0-28-lowlatency.

The computers are interconnected via Gigabit Ethernet. The SDR device is an NI USRP 2944R [12] which is connected via a ten Gigabit Ethernet connection to the computer that is running the FR-A with the embedded eNB implementation.

The ten Gigabit Ethernet connection is required in order to guarantee sufficient stability for the transfer of baseband I/Q samples between eNB and USRP in both uplink and downlink direction. This directly reflects the challenges faced by fronthaul implementations for lower layer splits in C-RAN architectures as discussed in detail in [18].

III. EXTENDED FLEXRAN COMMUNICATION

The communication between FR-C and FR-A has been extended in both directions in order to support adaptation of the resource scheduling by means of flexible bandwidth part adaptation and evaluation of computational resource utilization in terms of processing time requirements of individual physical layer functions. Corresponding messages and handlers have been implemented in both FR-C and FR-A.

---

[1] August 2017

A. *Bandwidth Part Adaptation*

The possibility to configure and adapt bandwidth parts (BWPs) within the system bandwidth is an essential feature of the new 5G NR air interface. According to [14], a BWP is a set of contiguous physical resource blocks (PRBs) selected from a contiguous subset of the common PRBs with a size that is lower than the total carrier bandwidth.

In order to implement flexible BWPs within the OAI framework for LTE, the interface between FR-A and FR-C has been extended so that the controller can configure the BWPs for the agent during runtime.

In order to comply with the resource block group (RBG) concept of LTE according to resource allocation type 0 in the specification [15], the BWP configuration and adaptation has been implemented on RBG level. In case of 10 MHz system bandwidth which has been used for the following performance study, an RBG comprises three consecutive PRBs.

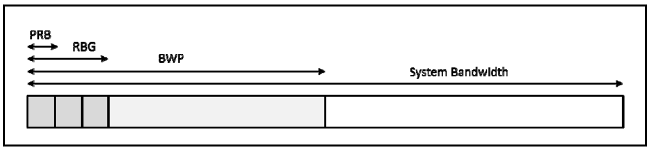

Fig. 2. Frequency domain resource allocation granularity

The relation between PRB, RBG, BWP and system bandwidth is shown in Fig. 2. It is important to keep in mind during the following performance study that the RBG determines the resource allocation granularity during downlink scheduling while the BWP determines the set of consecutive RBGs that are used for downlink scheduling.

A typical spectrum allocation snapshot is shown in Fig. 3 where a BWP comprising 12 PRBs is used for downlink resource scheduling. The figure furthermore shows the allocation of the center six PRBs of the system bandwidth every five milliseconds for primary and secondary synchronization signal (PSS and SSS) as specified for LTE. The remaining periodic power allocations within the whole system bandwidth mainly correspond to common reference symbols (CRS) and physical downlink control channels (PDCCHs) that are distributed over the whole system bandwidth.

Here it has to be kept in mind that the bandwidth part implementation used in this paper is not the exact realization of 5G NR bandwidth parts since the latter would for example contain additional reference and synchronization signals that cannot be used in combination with conventional LTE mobiles as done in Section IV.

In the initial performance study presented in this paper, the FR-C is basically responsible for assigning BWPs consisting of consecutive RBGs to the FR-A which hosts the OAI eNB implementation. This BWP allocation can be adapted during runtime of the eNB. The BWP concept implementation and the corresponding adaptation are both transparent for the downlink resource scheduler. The latter just operates on a set of PRBs that are available for downlink resource allocations within a transmission time interval (TTI). The BWP concept is in this sense from scheduler point of view therefore basically a PRB set restriction which makes it possible to use conventional LTE mobiles during the experiments.

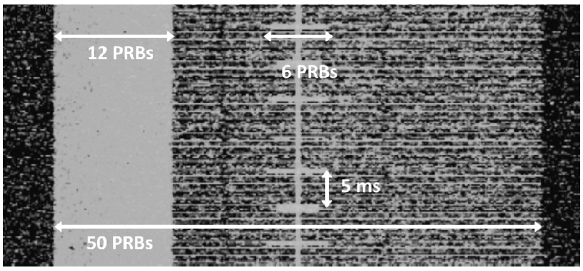

Fig. 3. Exemplary spectrum analyzer snapshot with 50 PRBs system bandwidth and a BWP comprising 12 PRBs

B. *Processing Time Reporting*

The time required for performing different PHY functions within the eNB are reported periodically every TTI, corresponding to an LTE subframe of one millisecond duration, from the FR-A to the FR-C. This high frequency of reporting is used in the study in order to facilitate an accurate estimation of processing time distributions.

The processing time reporting in the context of this paper comprises downlink resource scheduling, transport block encoding, transport block scrambling, and transport block modulation within the downlink transmission chain of the eNB. The three functions are called sequentially per scheduled UE during the TTI construction.

Spatial precoding is not considered in this paper since a system configuration with a single TX antenna port is used, meaning that no spatial multiplexing or transmit diversity schemes will be applied. This corresponds to LTE transmission mode 1 as specified in [15].

IV. PERFORMANCE STUDY

The performance measurements were conducted with different commercial LTE mobiles (Google Nexus, Samsung Galaxy S7, Samsung Galaxy Tab active) operating in E-UTRA FDD Band 7 (2.66 GHz downlink, 2.54 GHz uplink). The system bandwidth has been set to 50 PRBs, corresponding to a nominal channel bandwidth of 10 MHz. The downlink bandwidth part (BWP) configuration has been varied between 50 PRBs and 6 PRBs. All mobiles are scheduled within a single configured BWP.

Since the maximum downlink throughput performance is addressed in this paper, the mobiles have been positioned in a way that provides sufficient SINR levels (> 20 dB) for the highest modulation and coding scheme (MCS) level that is supported by the mobiles and by the OAI implementation.

A. *Downlink Throughput*

The downlink throughput has been evaluated depending on the number of active mobiles and the number of PRBs available for resource allocation in a configured BWP. The throughput on the physical layer has been evaluated on TTI level using the Accuver tool XCAL-M [17] which provides direct access to the user equipment (UE) chipsets.

The traffic load has been generated by using iPerf3 [16] between the S-GW and the UEs with an appropriate configuration for full buffer UDP traffic in order to achieve maximum throughput.

Table 1. Maximum theoretical downlink throughput (in Mbps)

		Number of PBRs				
		6	18	27	39	50
Number of UEs	1	4.4	13.5	19.8	29.3	36.7
	2	2.2	6.8	9.9	14.6	18.3
	3	1.5	4.5	6.6	9.8	12.2

The content of Table 1 provides the maximum achievable downlink throughput (in Mbit/s) on the LTE PHY. The throughput has been determined under the assumption of fair resource sharing between the UEs, meaning that every UE will get in average the same downlink throughput, and the assumption that the transport block size (TBS) is always chosen according to the highest MCS which corresponds to 64QAM as modulation scheme (assuming no support for 256QAM) as specified in [15].

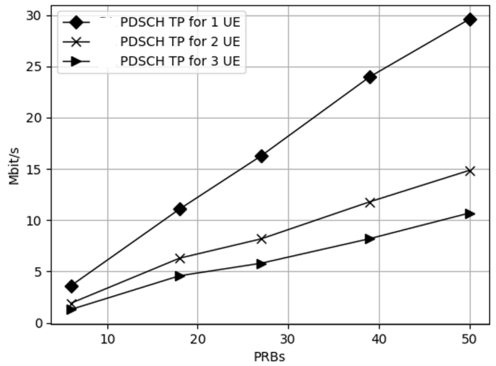

Fig. 4. Downlink throughput per UE depending on bandwidth part configuration

The result of the downlink throughput evaluation with OAI and conventional LTE mobiles (UEs) is shown in Fig. 4. The results show that the throughput grows linearly with the number of PRBs and that the resources are shared in fair fashion between all active UEs. The difference between theoretical throughput limit and the actual observed downlink throughput corresponds to approximately 20% loss. The reason for this is given by the fact that the OAI downlink resource scheduler that has been used does not schedule subframes for downlink transmissions that carry PSS and SSS. Since these are transmitted every fifth subframes, it yields a user throughput performance degradation of 20%.

B. Processing Time

The processing time is evaluated on TTI level, meaning that the time required for constructing the downlink transmission is measured in the FR-A (representing the DU) and reported to the FR-C (representing the CU) every TTI. Since the construction of a downlink subframe has to be done every millisecond in an LTE eNB, the overall processing time should never exceed this limit. The processing time requirement will become even more demanding under consideration of fronthaul latencies between CU and DU as discussed in [18]. The 5G NR air interface will furthermore support flexible TTI durations down to the scheduling of mini slot with 0.1 millisecond duration.

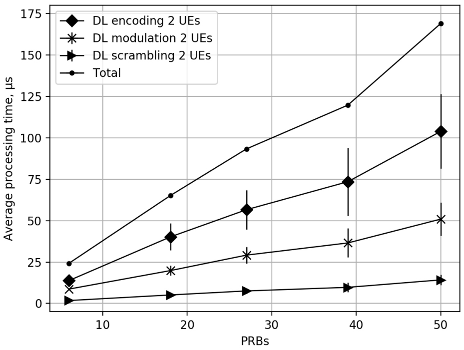

Fig. 5. Processing time evaluation with two active UEs

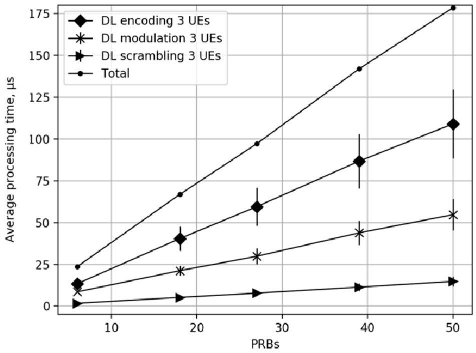

Fig. 6. Processing time evaluation with three active UEs

Fig. 5 and Fig. 6 show the processing time with two and three UEs, respectively. In addition to the average processing time per TTI, the figures show the standard deviation of the processing time for encoding, scrambling and modulation. It can clearly be seen that the processing time required for encoding exhibits the largest variance in comparison with modulation and scrambling. Especially the latter shows a negligible variance of the processing time. All throughput performance evaluations were running for 20 seconds, corresponding to 20.000 TTIs of 1 millisecond duration in order to collect a sufficiently large number of processing time reporting samples.

The processing time required for the downlink resource scheduling itself has been measured as well, but is not considered in the discussion here due to its observed limited impact on the overall downlink TTI processing time. The time required to perform the downlink resource allocation for a TTI was in all cases (one, two, and three active UEs) below 15 microseconds. The IFFT processing in the OFDM transmit path requires in the evaluated scenario with 10 MHz system bandwidth additionally 55 microseconds with negligible variance per TTI, independent of number of UEs and number of PRBs.

Especially the results with three UEs show that the required processing time for all three evaluated functional entities depends linearly on the number of PRBs. It is therefore possible to construct models for the processing time depending on PRB allocations for example by means of linear regression or polynomial approximation. Such models can then be used for the design of computational elasticity algorithms that adapt transmission parameters depending on computation resource utilization and availability.

Here it has to be taken into account that the results presented in this paper are just exemplary initial measurements since the overall processing time in specific configurations will always depends on implementation details and on the CPU clock rate.

A more detailed analysis of the sum processing time distribution comprising encoding, scrambling and modulation is given in Table 2. It provides the probabilities for exceeding certain processing time values depending on the number of active UEs with downlink traffic. The results indicate that the processing time variance depends much more on the number of UEs than the average processing time. This has to be taken into account when designing future computational elasticity strategies.

Table 2. Detailed information on processing time distribution depending on number of UEs

		Number of UEs		
		1	2	3
Processing time	> 200 µs	0.05	0.39	0.83
	> 205 µs	0.00	0.01	0.27
	> 210 µs	0.00	0.00	0.01

V. CONCLUSION

In this paper, we presented an extension of the OpenAirInterface based FlexRAN interface between controller and agent corresponding to splitting the LTE eNB functionality between centralized unit (CU) and distributed unit (DU) within a cloud based RAN. The interface extension focuses on the adaptation of bandwidth part allocations and on periodic processing time reporting of individual physical layer functions for the facilitation of computational elasticity strategies.

Initial performance evaluations revealed that the processing time of the considered physical layer functions increases linearly with the number of bandwidth (in terms of PRBs) used for downlink transmissions. The statistical evaluation has furthermore revealed that the encoding process shows significantly larger processing time variance than modulation and scrambling.

Next steps will focus on the development of specific control algorithms for computational elasticity based on the presented interface and further enhancements of the interface taking into account dynamic adaptation of MCS level and spatial precoding.

ACKNOWLEDGMENT

This work has been performed in the framework of the H2020 project 5G-MoNArch co-funded by the EU. The views expressed are those of the authors and do not necessarily represent the project. The consortium is not liable for any use that may be made of any of the information contained therein.

REFERENCES

[1] ITU-R M.2083-0, “IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond”, Sep. 2015

[2] NGMN 5G Initiative White Paper, Feb. 2015

[3] 3GPP Technical Specification Group Radio Access Network, TR 38.801 v14.0.0, “Study on new radio access technology: Radio architecture and interfaces”, March 2017

[4] A. Checko, H. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. Berger, and L. Dittmann “Cloud RAN for Mobile Networks – A Technology Overview”, IEEE Comm. Survey & Tutorials, vol. 17, no. 1, 2015

[5] P. Rost, C. Bernardos, A. De Domenico, M. Di Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, “Cloud Technologies for Flexible 5G Radio Access Networks”, IEEE Communications Magazine, vol. 52, no. 5, May 2014

[6] 3GPP Technical Specification Group Radio Access Network, TS 38.300 v15.0.0, “NR and NG-RAN Overall Description”, Dec. 2017

[7] P.Arnold, N. Bayer, J. Belschner, and G. Zimmermann, “5G Radio Access Network Architecture Based on Flexible Functional Control / User Plane Splits”, 2017 European Conference on Networks and Communications , June 2017

[8] Open Air Interface, www.openairinterface.org

[9] X. Foukas, N. Nikalein, M. Kassem, M. Marina, and K. Kontovasilis, “FlexRAN: A Flexible and Programmable Platform for Software-Defined Radio Access Networks”, 12th International Conference on emerging Networking EXperiments and Technologies , Dec. 2016

[10] M. Jammal, T. Singh, A. Shami, R. Asal, and Y. Li, “Software defined networking: State of the art and research challenges”, Comptuter Networks, vol. 72, Oct 2014

[11] Protocol Buffers, https://developers.google.com/protocol-buffers/

[12] Nation Instruments, www.ni.com

[13] I.A. Alimi, A.L. Teixeira, P.P. Monteiro, “Toward an Efficient C-RAN Optical Fronthaul for the Future Networks: A Tutorial on Technologies, Requirements, Challenges, and Solutions”, IEEE Comm. Survey & Tutorials, vol. 20, no. 1, 2018

[14] 3GPP Technical Specification Group Radio Access Network, TS 38.211 v15.0.0, “NR; Physical channls and modulation”, Dec. 2017

[15] 3GPP Technical Specification Group Radio Access Network, TS 36.213 v15.1.0, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures”, March. 2018

[16] iPerf3, https://iperf.fr/

[17] Accuver, www.accuver.com

[18] I. Alyafawi, E. Schiller, T. Braun, D. Dimitrova, A. Gomes, and N. Nikaein, “Critical Issues of Centralized and Cloudified LTE-FDD Radio Access Networks”, 2015 IEEE International Conference on Communications , June 2015