

Research Article

Fronthaul for Cloud-RAN Enabling Network Slicing in 5G Mobile Networks

Line M. P. Larsen , **Michael S. Berger**, and **Henrik L. Christiansen**

Department of Photonics Engineering, Technical University of Denmark, 2800, Denmark

Correspondence should be addressed to Line M. P. Larsen; lmph@fotonik.dtu.dk

Received 20 April 2018; Revised 2 July 2018; Accepted 17 July 2018; Published 28 August 2018

Academic Editor: Pei Xiao

Copyright © 2018 Line M. P. Larsen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work considers how network slicing can use the network architecture Cloud-Radio Access Network (C-RAN) as an enabler for the required prerequisite network virtualization. Specifically this work looks at a segment of the C-RAN architecture called the fronthaul network. The fronthaul network required for network slicing needs to be able to dynamically assign capacity where it is needed. Deploying a fronthaul network faces a trade-off between fronthaul bitrate, flexibility, and complexity of the local equipment close to the user. This work relates the challenges currently faced in C-RAN research to the network requirements in network slicing. It also shows how using a packet-switched fronthaul for network slicing will bring great advantages and enable the use of different functional splits, while the price to pay is a minor decrease in fronthaul length due to latency constraints.

1. Introduction

Emerging technologies paving the way towards the next generation, 5G, mobile networks include the very promising concept of network slicing. Network slicing describes how one physical network is divided into multiple logical networks, referred to as network slices. One network slice can have specific capabilities related to one specific service, like one slice for Internet of Things (IoT), as this service has specific requirements to the network. Another benefit of network slicing is that it is possible for several operators to share the same physical network and thereby save cost for deployment and maintenance of the physical network equipment. Different network slices have different requirements to the network, because different applications have different requirements to the network [1]. The 3rd Generation Partnership Project (3GPP) describes in [2] how network slicing is envisioned to provide different capabilities on different slices. These capabilities can be related to the three primary 5G drivers [3]:

- (i) Extreme Mobile Broadband (eMBB) will support use cases like “shopping mall” requiring 300 Mbps experienced Downlink (DL) throughput and at least 95% availability and reliability for all applications [3].

- (ii) Massive Machine-Type Communication (mMTC) will support use cases like “massive amount of geographically spread devices” requiring up to 1,000,000 devices per km² and 10 years of battery life [3].
- (iii) Ultrareliable Machine-Type Communication (uMTC) will support use cases like “autonomous vehicle control” requiring latency below 5 ms and 99.999% availability and reliability [3].

These examples show how different the requirements can be, to different slices within network slicing. The same physical resources need to be able to carry very different demands, which do not only require complex QoS management but also put huge requirements to the physical network that should be able to handle it. Network slicing requires a virtualization of the network to be able to run several logical networks on top of the physical network [4].

Cloud-Radio Access Network (C-RAN) is a promising network architecture which can be used to enable virtualized networks and network slicing. In C-RAN the base station functions known from the protocol stack are divided into a Distributed Unit (DU) and a Centralized Unit (CU). The DU is located close to the antenna in the antenna mast and is thereby close to the user, where the CU can be located in a datacenter benefitting from high processing powers.

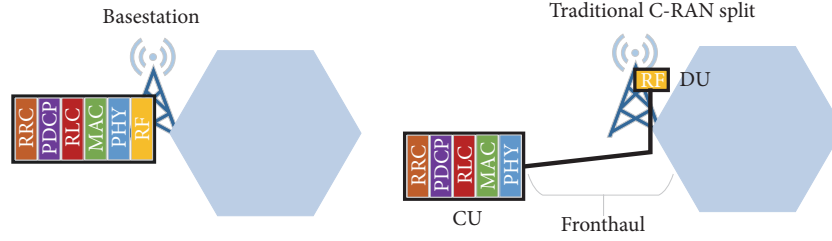


FIGURE 1: Function allocation in a base station and in the traditional C-RAN split.

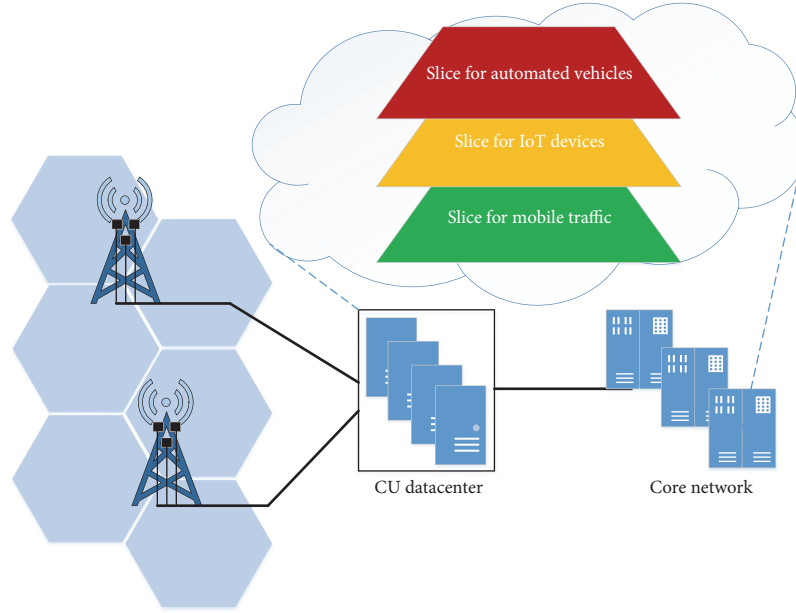


FIGURE 2: Network slicing in Cloud-RAN.

The exact location of the separation between these two entities is referred to as the functional split. The DU and the CU are connected using a so-called fronthaul network. The simplest division between DU and CU leaves only the Radio Frequency (RF) functions in the DU. This division will be referred to as the traditional split throughout this paper and the division of functions in the traditional split is illustrated in Figure 1.

Figure 1 illustrates the traditional split where the base station functions are illustrated using the LTE protocol stack, as no protocol stack exists for 5G at the time of writing. In the traditional split, raw in-phase quadrature (IQ) samples are transported on the fronthaul link, resulting in a very high and constant bitrate. Using the traditional split, the IQ samples need a special protocol for transport over the fronthaul network. Several options exist including the widely used Common Public Radio Interface (CPRI) [6]. Originally, CPRI was intended for point-to-point transmission, but this makes the fronthaul link very inflexible, as each CU/DU pair requires its own fiber connection. Seen from a network slicing perspective, this solution is not very flexible, as each slice gets a static amount of capacity assigned for fronthaul transmission.

CUs from several sites can be centralized in the same datacenter which is an enabler for modern network virtualization techniques. This way, processing functions are gathered in one place, the CU-datacenter, which can be virtualized. Network functions virtualization moves the network processes into software, and, instead of the functions running at a base station, they will be able to run at any server [7]. Virtualization of several functions is an important enabler for network slicing. The situation is illustrated in Figure 2.

Figure 2 illustrates how network virtualization of the CU-datacenter is used to run several logical network slices on top of one physical network. The logical slices serve different purposes as each of them complies with different network requirements. Virtualizing C-RAN brings benefits in scalable management of processing resources and enables network programmability [8]. This work looks into how network slicing can be enabled using C-RAN with a special focus on the issues raised in future fronthaul networks. The remainder of this paper contains an overview of the current trends being investigated within fronthaul deployment, a comprehensive description of the fronthaul challenges that are still under research, and a discussion of the options existing for C-RAN deployment.

TABLE 1: Comparison of a circuit switched and a packet switched fronthaul link.

	Circuit switched fronthaul	Packet switched fronthaul
Pros	Guaranteed bitrate.	Adaptable to non-uniform traffic.
Cons	Load independent of cell load.	Delay can occur
Capacity	Guaranteed	Depends on technology
Timing	Guaranteed	Depends on technology
Synchronization	Guaranteed	Depends on technology
QoS	Dedicated user channel	Shared transmission
Service Guarantee	Dedicated resources	Delay can occur
Multiplexing	TDM or WDM	Statistical multiplexing
Resource utilization	Constant use of resources	Improved

2. Recent Trends in Mobile Fronthaul

The mobile fronthaul network has been considered in several papers, investigating different solutions and options. C-RAN and network slicing are two concepts that have already been combined in several works. Reference [5] contributes with a dynamic network slicing scheme for multitenant C-RANs. In [9] a demo of network slicing using C-RAN is introduced which aims at efficiently sharing the bandwidth resources among different slices.

A large survey on cloud-based services in [8] states the benefits of virtualized networks and argues for the use of the C-RAN architecture due to its scalability and high system capacity. The standardization of C-RAN and the division of functions between the CU and DU are a currently on-going process, where 3GPP contributes with [10] and IEEE have established the 1914 Next Generation Fronthaul Interface (NGFI) working group [11]. Also different industry alliances are looking into the subject including the CPRI consortium [12], NGMN alliance [13], and Small Cell Forum [14]. In [15], the authors argue for using the Ethernet network as the new fronthaul, as it is already deployed and brings several benefits. Reference [16] reports on the performance of different functional splits over a bridged Ethernet network; results show fronthaul processing delays and how these are affected by different types of traffic flows. Different papers and organizations use different naming or numbering of the functional splits. To illustrate the functional splits from a more generic point of view, this paper will not refer to any names, but only to the locations of the functional splits in the LTE protocol stack. And only those functional split locations with specific fronthaul abilities will be mentioned, as many more exist. For a complete overview of the options currently researched are referred to in [17].

3. Fronthaul Link Type

The performance of a fronthaul network depends on whether a circuit-switched solution or a packet-switched solution is chosen. The capabilities of a circuit- and a packet-switched fronthaul network are summarized in Table 1.

Circuit-switched solutions, for example, Optical Transport Network (OTN), provide a constant use of resources, as the connections are always using the same amount of capacity, which is statically assigned. This means that queuing

will be very limited as the resources are always reserved and the network provides a stable connection in terms of capacity, timing, and synchronization. OTN is a circuit-switched solution that provides protection and multiplexes several transmissions using WDM [18]. The pros of using OTN are that it uses Forward Error Correction (FEC) and it can measure the Bit Error Rate (BER). But when using OTN, or another circuit-switched technology in the fronthaul network, the capacity is fixed to the already assigned carriers and is not able to assign extra capacity to establish or scale slices dynamically.

In packet-switched technologies, it is possible to dynamically assign capacity when needed. Using statistical multiplexing, a packet-switched network is adaptable for a variable load on the fronthaul, because multiple connections are multiplexed into one fiber. Ethernet is an example of a packet-switched fronthaul technology. Ethernet allows sharing of the network infrastructure through standardized virtualization techniques and, through its packet-switched operation, it will save resources on the fronthaul link using statistical multiplexing gain [15]. This makes Ethernet able to flexibly allocate resources and, when considering network slicing, dynamically reserve resources for specific slices. Ethernet has several advantages to be used for fronthaul; it is flexible, cost-effective, and widely used. The Ethernet network has one problem though, before it can be used as the new fronthaul network; in a packet-switched technology, traffic is more eager to queue. Queuing traffic can create unwanted delay and jitter in a mobile network, and therefore the industry is looking into solutions to avoid that using for example carrier grade management. In Ethernet the latency will also be affected by the number of switches to be passed through; this is the reason why Time Sensitive Networking (TSN) is providing options for traffic prioritization in the standard for frame preemption [19].

The trend is pointing towards using Ethernet as the fronthaul network [15], and this will be the focus in the remainder of this paper.

4. Challenges in Mobile Fronthaul

The introduction of network slicing requires a virtualization of the network and an efficient use of resources in the

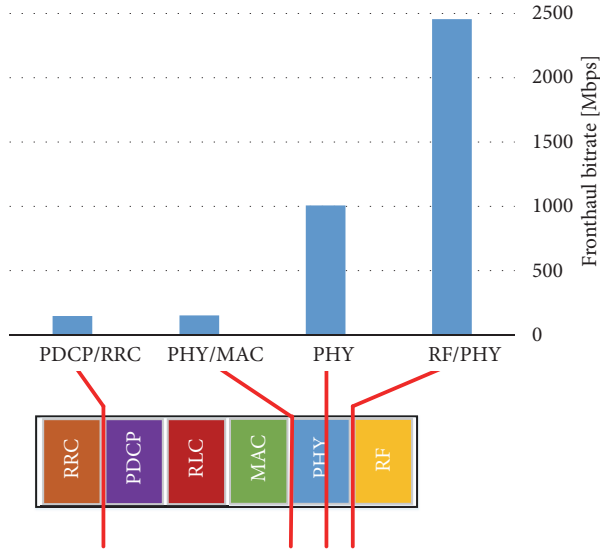


FIGURE 3: Fronthaul bitrates for selected functional splits. Note that, as per formulas in [5], the CPRI linecode is included in the bitrate for “RF/PHY” split, but not in the “PHY” split.

fronthaul network. This chapter looks into how Ethernet can be the solution to that.

4.1. Functional Split. A packet-switched technology, like Ethernet, does not obtain many benefits when transmitting data requiring a constant bitrate. Therefore, with the user bitrates highly increasing in 5G [2], the traditional split is not a sustainable solution due to the very high and constant fronthaul bitrate. The term functional split defines different options for splitting up the functions in the protocol stack. The trend points towards including more functions in the DU compared to the traditional split. When only a few functions are left in the DU, the signal is raw and the fronthaul bitrate is high with a constant load. Adding more functions in the CU will decrease the fronthaul bitrate and increase the fronthaul flexibility, as the signal gets more processed before the transmission, resulting in a fronthaul bitrate that will vary with the user load. But the low and variable bitrate comes with a cost; the DUs become more complex and thereby more difficult to install and maintain. A variable bitrate on the fronthaul network is crucial for Ethernet to perform dynamical resource allocation, and the lower the bitrate is, the more the resources can be multiplexed into the same fiber. The fronthaul bitrate can be calculated by looking at what type of data is actually being transmitted on the fronthaul link, which is different depending on what functional split is chosen. To state an example of the huge difference between the different functional splits, the fronthaul bitrates for different functional split options are illustrated in Figure 3. The fronthaul bitrates are calculated based on formulas from [14] and considering a 20 MHz LTE carrier using two antenna ports and 64 QAM modulation. The fronthaul bitrates are calculated for four different splits: the traditional split between the RF and physical layer (RF/PHY),

a split in the physical layer (PHY), a split between the physical layer and the Media Access Control (PHY/MAC), and a split between the Packet Data Convergence Protocol and Radio Resource Control (PDCP/RRC). The bitrates and the corresponding split location in the LTE protocol stack. The splits and their corresponding fronthaul bitrates are illustrated in Figure 3. Here functions on the left side of the red line are included in the CU and functions on the right side are included in the DU.

Figure 3 illustrates the locations of four different functional split options, and their corresponding fronthaul bitrates. The bitrate for the RF/PHY option is very high, and the bitrates for the PDCP/RRC and PHY/MAC options are very low. The functional split will also determine whether the bitrate on the fronthaul link is constant or varies with user load. This is determined by the amount of functions left in the DU, i.e., the amount of signal processing from the physical layer taking place in the DU. Looking at Figure 3, then the RF/PHY split has a constant load, where the PHY/MAC, PDCP/RRC splits have bitrates varying with user load on the fronthaul link. Whether the fronthaul bitrate is variable or not for the PHY split is depending on where in the physical layer the PHY split is located. Figure 4 illustrates the functional blocks within the physical layer and what types of data is transported in-between them. The blocks that affect the fronthaul bitrate the most are highlighted and will be further discussed; the remaining blocks are included in the figure for completeness. Read from the right side, data is received/transmitted from the RF block represented by IQ symbols. The red line illustrates how functional splits that includes the RE (de)mapper in the DU have a variable load on the fronthaul link, where functional splits located closer to the RF block have a constant load on the fronthaul link.

Read from the right towards left, Figure 4 illustrates how the DL signal received by the antennas are transmitted via the antenna ports into processing in the physical layer before it is further sent to the MAC using transport blocks, and reverse for the receiver. In the RF block, the signal is up/down converted and sampled. When entering the PHY block, the cyclic prefix is removed and sent into the Fast Fourier Transformation (FFT) where the signal is converted into subcarriers in the frequency domain. The FFT process induces a large reduction on the fronthaul bitrate, which in LTE corresponds to 40%, due to removal of guard subcarriers [20]. The Resource Element (RE) mapper maps between subcarriers and symbols. In this process, the unused subcarriers forwarded from the FFT can be detected. This way the RE mapper makes the signal vary with the user load [20]. Therefore, DUs that include the RE mapper will have a variable bitrate on the fronthaul link. And DUs that do not include the RE mapper will have a constant bitrate on the fronthaul link; i.e., functional splits on the left side of the red line have variable bitrate, and functional splits on the right side have constant bitrate on the fronthaul link. Another process in PHY that will further reduce the fronthaul bitrate is the modulation. When modulating the signal, a certain amount of bits will be mapped into one symbol, depending on the order of modulation [21]. The order of modulation will then determine how much the signal is reduced.

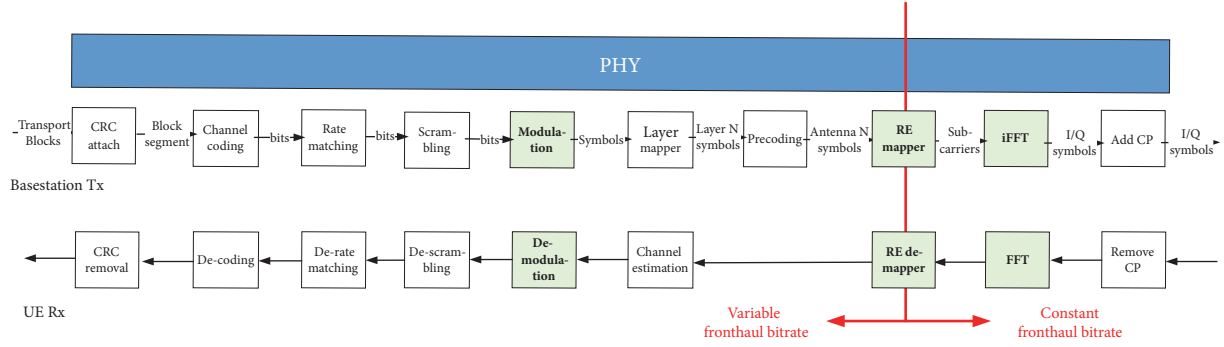


FIGURE 4: Functional blocks in the physical layer, shown for both uplink and downlink direction.

Using transport blocks the signal is transmitted from the physical layer into the MAC. In the LTE MAC, the Hybrid Automatic Repeat Request (HARQ) process is located together with a scheduler determining the cooperation among DUs [10]. The HARQ process is very time critical [21] and this puts large latency constraints on the fronthaul network when this process is located in the CU-datacenter. The scheduler on the other hand is beneficial to have in the centralized CU-datacenter for better management of the DU resources. Some functional splits are proposed to separate the scheduler into a local scheduler block in the DUs and a central scheduler in the CU-datacenter [10].

4.2. Fronthaul Delay. It is crucial for the fronthaul link to comply with different delay requirements, both because different functions within the protocol stack have different delay requirements and because different applications have different requirements to the response time. In network slicing, where all services rely on one network, the network needs to be compatible with the worst case situations. Using a packet-switched network, an extra delay can be expected due to queuing or a slight delay added in frame preemption for the high priority packets [19]. The delay on the fronthaul link is crucial to determine the length of the fronthaul network, as the distance between a DU and the CU-datacenter is determined by the maximum delay. Since certain functions and services determine the delay limits in the network, this delay will also determine the maximum distance between the DU and the CU-datacenter. The delay is given in

$$\text{Delay} = 2 \cdot \text{transmission delay} + \text{processing delay} \quad (1)$$

The transmission delay is calculated as

$$\text{Transmission delay} = \frac{\text{Packet size}}{\text{fronthaul bitrate}} \quad (2)$$

The maximum Ethernet packet size is 1500 bytes. The fronthaul bitrates are calculated based on formulas in [14].

The Round Trip Time (RTT) describes the time from a request is sent until a message is received, and therefore it also includes the propagation delay, i.e., the time for the request to propagate through the given medium with a certain length. The RTT is given in

$$\text{RTT} = \text{delay} + 2 \cdot \text{propagation delay} \quad (3)$$

The RTT delay needs to be compliant with the delay requirements for the specific service that it is running, and when an Ethernet fronthaul is assumed, the delay for queuing and processing through Ethernet needs to be considered. Therefore a delay for one Ethernet switch is added as D_{sw} and multiplied by the number of switches:

$$\text{RTT} > \text{delay} + 2 \cdot D_{sw} \cdot \# \text{ switches} + 2 \cdot \text{propagation delay} \quad (4)$$

The delay in one switch, D_{sw} is depending on the type of switch and the packet size. The following equation is an assumption for D_{sw} assuming a Gb Ethernet switch is used:

$$D_{sw} = \frac{\text{Packet size}}{1 \cdot 10^9} \cdot 1000 \text{ ms} \quad (5)$$

$\cdot \text{switch processing time and queuing delay}$

The total switch processing time and queuing delay can be adjusted according to amount of traffic and priority packets. In this work, it is assumed to be factor 3.

The propagation delay is calculated as

$$\text{Propagation delay} = \text{propagation delay per km} \cdot \text{fronthaul range} \quad (6)$$

The distance between the DU and CU, the fronthaul range, can be determined by

$$\text{Fronthaul range} \leq \frac{(\text{RTT} - \text{delay})}{\text{propagation delay per km}} \quad (7)$$

As an example, eMBB is considered as the service giving the delay requirements, here the maximum latency for the user plane is 4 ms [22], corresponding to 8 ms RTT. The fiber propagation delay is 10 $\mu\text{s}/\text{km}$ [23].

Figure 5 illustrates how much the fronthaul range is affected when the transmitted data has to pass through different numbers of switches. Depending on the processing delay, the fronthaul range can be several hundred of km. It must be considered that the processing time might be lower in the CU-datacenter compared to the standalone DU. Therefore the processing delay has to be distinguished

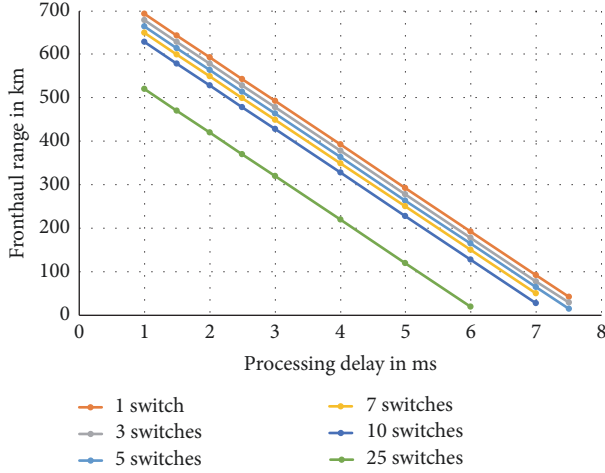


FIGURE 5: Fronthaul range corresponding to different amounts of fronthaul delay illustrated for different numbers of switches in the fronthaul network.

between the amount of functions being processed in the DU and in the CU-datacenter. The amount of switches the data needs to pass on the way between the DU and the CU-datacenter will most likely depend on the population density in the current area, as a higher population density might incur more switches in the area. Looking at Figure 5, then the difference between passing one switch and 25 switches gives a difference in the fronthaul range of approximately 170 km.

Some functions within the LTE protocol stack have higher requirements to the max delay; this is for example the HARQ process located in the MAC, which is limited by a RTT of 5 ms [21]. And because LTE might need to coexist with other RATs on its own network slice, HARQ is also considered here. In splits where the HARQ process is located in the CU-datacenter, the signal needs to be transferred over the fronthaul network and back within the 5 ms RTT. The RTT of the HARQ located in the DU and the CU, respectively, is illustrated in Figure 6.

Figure 6 shows how the distance to overcome when the HARQ process is located in either the CU or the DU. When the HARQ process is located in the DU, the distance to the user is very short and the signal only has to travel from the user to the DU and back within the 5 ms of RTT. Therefore, the latency for the splits where the HARQ process is included in the DU is more relaxed and results in a much longer fronthaul range. The location of the MAC and HARQ process in the LTE protocol stack is illustrated in Figure 7.

Due to the HARQ process's delay limitations, three delay scenarios exist: one for the splits before the RE mapper, having a constant load on the fronthaul link, one for the remaining splits where the HARQ is located in the CU-datacenter, and one for the splits where the HARQ is located in the DU and the latency is limited by a certain service. The latter one has already been described.

The delay budget for the splits before the RE mapper, with no Ethernet delay added as the connection is expected to be circuit-switched:

$$5 \text{ ms} > \text{delay} + 2 \cdot \text{propagation delay} \quad (8)$$

The delay budget for the split options after the RE mapper, where the HARQ is still located in the CU, here an Ethernet fronthaul is assumed and therefore a delay for one Ethernet switch is added as D_{sw} and multiplied by the number of switches:

$$5 \text{ ms} > \text{delay} + 2 \cdot D_{sw} \cdot \# \text{ switches} + 2 \cdot \text{propagation delay} \quad (9)$$

Figure 8 illustrates the fronthaul range limited by the HARQ process. The number of switches is 5.

Figure 8 shows how the HARQ requirements affect the range of the fronthaul link. The figure also shows how different processing delays will affect the fronthaul range. Applications that have larger requirements to the delay need to follow this limitation if they use the traditional functional split or a split where the HARQ is located in the CU. Applications that have the HARQ located in the DU have larger requirements to the delay resulting in a much longer fronthaul range illustrated in Figure 5. Figure 8 clearly shows that the delay added by the Ethernet network affects the fronthaul length, as the options “Before RE map” and “HARQ in CU” have a clearly division. Theoretically, the processing delay should be lower for the functions implemented in the CU-datacenter, as those have higher processing powers available compared to the single DUs at the cell sites. This would in that case affect the RTT, but this has not been taken into consideration in these calculations. Figure 8 shows how using an Ethernet network as fronthaul has a minor impact on the delay and thereby the fronthaul range.

5. Discussion

In a physical network used for network slicing, one network has to carry the traffic from several logical networks, each having very different requirements to the physical network. The one physical network needs to be able to run all extreme scenarios at the same time. It must have high capacity in terms of bitrates and number of supported devices and it must be extremely resilient and support ultralow latency, all at the same time for the network resources to be utilized on the right slices. C-RAN networks open up for the opportunity to share processing resources and allocate extra resources where they are needed as several functions are incorporated in a datacenter. But when using a circuit-switched fronthaul, resources are statically assigned and provide a dedicated transmission. Therefore a trade-off exists between latency added in packet queues and dynamical resource allocation in the fronthaul network. The trend points towards using the already established Ethernet network as fronthaul. If a solution using Ethernet or another packet-switched fronthaul technology is chosen, the capabilities provided by TSN will be highly necessary to prioritize the functions with very strict latency requirements. In Ethernet the physical resources can be utilized in a more efficient manner by flexibly allocating capacity within the fibers and the switches corresponding to a specific slice's demands. Using a flexible resource allocation, it is optional whether TSN is used in one slice or not. It is also optional what functional split is used in a specific slice

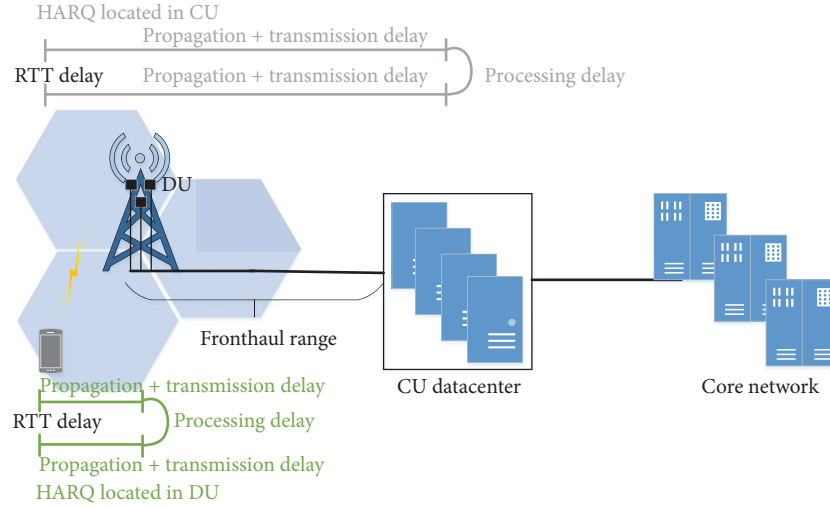


FIGURE 6: RTT response when HARQ is located in the CU-datascenter and when HARQ is located in the DU.

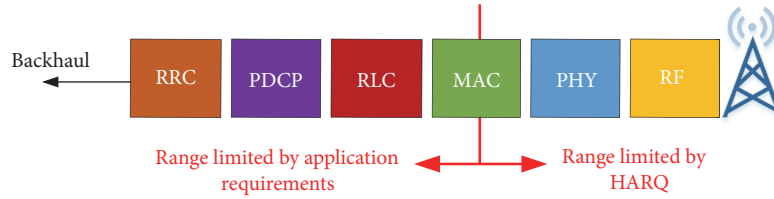


FIGURE 7: LTE protocol stack illustrating range limitations for splits before and after MAC/HARQ process.

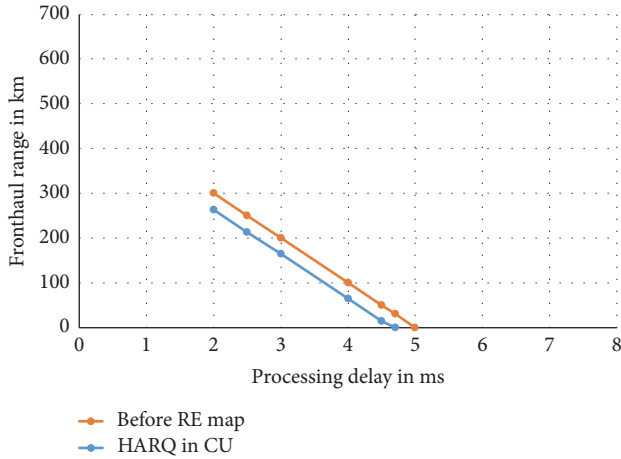


FIGURE 8: Fronthaul range for LTE scenarios complying with HARQ requirements.

and so on. One big downside of using a packet-switched network is usually that the delay is enhanced, but due to the calculations prior to Figure 8, the delay is minor. The length of the fronthaul, referring to Figures 5 and 8, is depending on the delay requirements and the available processing powers.

A crucial parameter is to decide which functional split to use, or in which situations a certain functional split shall be used. Different functional splits have different pros and cons, which is partly related to what benefits the different

functions have when they are local, close to the user or when they are centralized benefitting from large processing powers in the CU-datascenter. Hence different functional splits can be used in different scenarios. Some scenarios might require a large amount of centralized processing and a simple DU, where other scenarios obtain benefits in separating the user plane and control plane. Other functional splits are in-between these extremities: they want the benefits from a simplified DU but they also want a low fronthaul bitrate. Seen from a network slicing perspective it is recommended to use functional splits that have a variable fronthaul bitrate and runs over a packet-switched network. Considering the primary 5G drivers from [3], then eMBB will benefit from a variable bitrate on the fronthaul link; as enormous amounts of data needs to be transmitted at peak times, it will also benefit from a large amount of centralized processing, for a more efficient allocation of resources. mMTC has relaxed bitrate requirements; therefore it will benefit from a simple DU for easy deployment and maintenance. uMTC has very strict requirements to the latency, and thereby it requires a network with good traffic management and flexible resource allocation in the fronthaul network for faster transmission, and it will benefit from a centralized MAC scheduler. If C-RAN carries the network architecture for the physical network in network slicing, this one network needs to be compatible with all the requirements in all slices. In this manner it would make most sense to use different functional splits for different slices, in order to obtain the best resource utilization and performance.

6. Conclusion

C-RAN is a promising network architecture enabling virtualization techniques to be used for example for network slicing. In C-RAN the sites are connected, using a fronthaul network, to high capacity datacenters running more or less base station processing functions. Deploying a fronthaul network faces a trade-off between fronthaul bitrate, flexibility, and complexity of the local equipment close to the user. Network slicing introduces the concept of one physical network running several logical networks with different requirements on top. As the different slices can have very large and differentiated requirements to the network, the physical network needs to be equipped to handle extreme scenarios. This work showed how using a packet-switched fronthaul, for network slicing will bring great advantages and enable the use of different functional splits, while the price to pay is a relatively minor delay.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the Eurostars “Fronthaul for C-RAN” Project funded by Innovation Fund Denmark.

References

- [1] S. Sharma, R. Miller, and A. Francini, “A cloud-native approach to 5G network slicing,” *IEEE Communications Magazine*, vol. 55, no. 8, pp. 120–127, 2017.
- [2] 3GPP TS 22.261 V16.3.0, “Service requirements for the 5G system; Stage 1,” 2018.
- [3] A. Osseiran, J. F. Monserrat, and Y. P. Marsch, *5G Mobile and Wireless Communications Technology*, Cambridge University Press, Cambridge, UK, 2016.
- [4] T. Yoo, “Network slicing architecture for 5G network,” in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC '16)*, pp. 1010–1014, IEEE, Jeju Island, Korea, October 2016.
- [5] Y. L. Lee, J. Loo, T. C. Chuah, and L. Wang, “Dynamic network slicing for multitenant heterogeneous cloud radio access networks,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2146–2161, 2018.
- [6] CPRI Consortium, “CPRI Specification V7.0 Common Public Radio Interface (CPRI); Interface Specification,” 2015.
- [7] B. Chatras, U. S. Tsang Kwong, and N. Bihannic, “NFV enabling network slicing for 5G,” in *Proceedings of the 20th Conference on Innovations in Clouds, Internet and Networks (ICIN '17)*, pp. 219–225, March 2017.
- [8] G. P. Xavier and B. Kantarci, “A survey on the communication and network enablers for cloud-based services: state of the art, challenges, and opportunities,” *Annals of Telecommunications-Annales des Télécommunications*, vol. 73, no. 3-4, pp. 169–192, 2018.
- [9] S. Costanzo, I. Fajjari, N. Aitsaadi, and R. Langar, “DEMO: SDN-based network slicing in C-RAN,” in *Proceedings of the 15th IEEE Annual Consumer Communications & Networking Conference (CCNC '18)*, pp. 1-2, 2018.
- [10] 3GPP, “3GPP TR 38.801 V14.0.0 (2017-03): Study on new radio access technology: Radio access architecture and interfaces,” 3GPP, 2017.
- [11] IEEE 1914, <http://sites.ieee.org/sagroups-1914/>, Accessed 2018-13-02.
- [12] “eCPRI Specification V1.1 Common Public Radio Interface: eCPRI Interface Specification.”
- [13] NGMN, “NGMN Overview on 5G RAN Functional Decomposition,” 2018.
- [14] Small Cell Forum, “Solving the HetNet puzzle Small cell virtualization functional splits and use cases.”
- [15] N. J. Gomes, P. Sehier, H. Thomas et al., “Boosting 5G through ethernet: how evolved fronthaul can take next-generation mobile to the next level,” *IEEE Vehicular Technology Magazine*, vol. 13, no. 1, pp. 74–84, 2018.
- [16] P. Assimakopoulos, G. S. Birring, M. K. Al-Hares, and N. J. Gomes, “Ethernet-based fronthauling for cloud-radio access networks,” in *Proceedings of the 19th International Conference on Transparent Optical Networks (ICTON '17)*, pp. 1-4, July 2017.
- [17] L. M. P. Larsen, A. Checko, and H. L. Christiansen, “A survey of the functional splits proposed for 5G mobile crosshaul networks,” *IEEE Communications Surveys & Tutorials*, 2018.
- [18] Y. Ma, X. Huo, J. Li, Xiaomu, and Jingwen, “Optical solutions for fronthaul application (invited),” in *Proceedings of the 14th International Conference on Optical Communications and Networks (ICOON '15)*, pp. 1-3, July 2015.
- [19] IEEE computer Society, “IEEE Std 802.1Qbu: Frame Preemption,” 2016.
- [20] D. Wübben, P. Rost, J. S. Bartelt et al., “Benefits and impact of cloud computing on 5G signal processing: flexible centralization through cloud-RAN,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 35–44, 2014.
- [21] M. Sauter, *From GSM to LTE: An Introduction to Mobile Networks and Mobile Broadband*, 2011.
- [22] 3GPP, “TR 38.913 V14.3.0 (2017-06): Study on Scenarios and Requirements for Next Generation Access Technologies.”
- [23] D. Chitimalla, K. Kondepudi, L. Valcarenghi, M. Tornatore, and B. Mukherjee, “5G fronthaul-latency and jitter studies of CPRI over ethernet,” *Journal of Optical Communications and Networking*, vol. 9, no. 2, pp. 172–182, 2017.

