Cloud Radio Access Network architecture. Towards 5G mobile networks

Checko, Aleksandra

Publication date:
2016

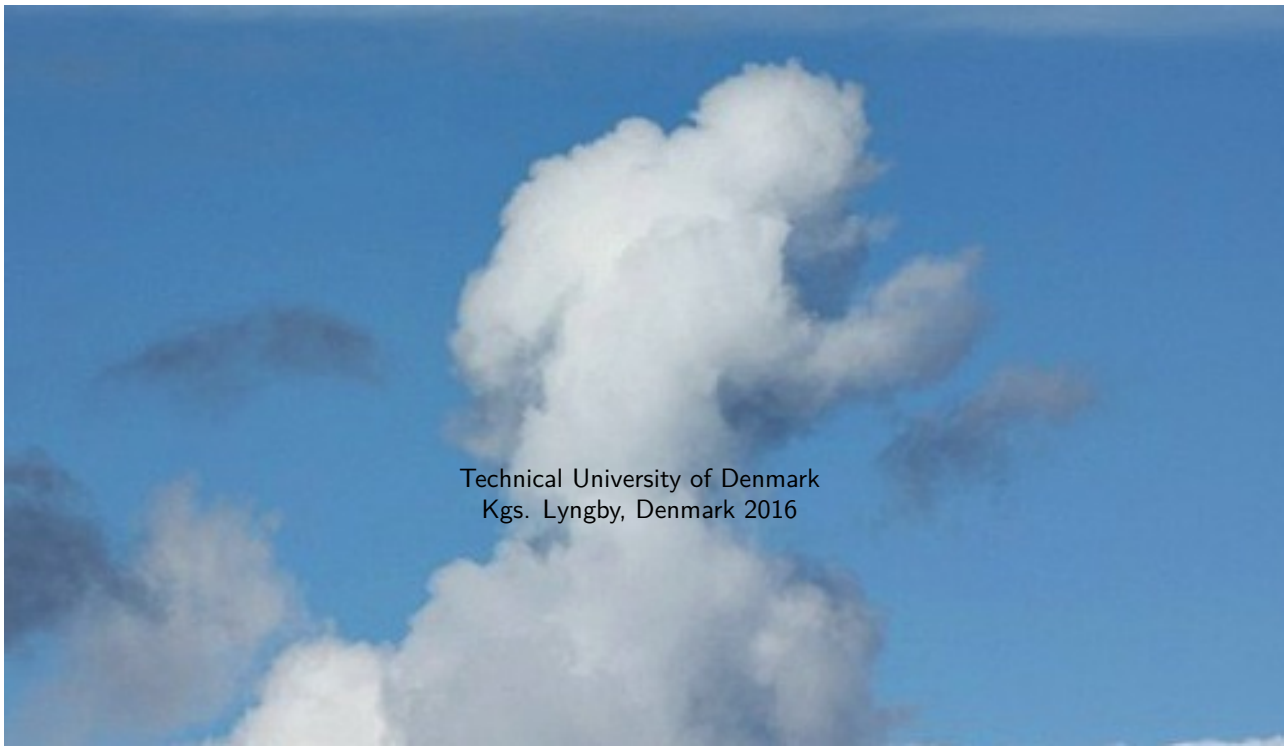
Document Version
Publisher's PDF, also known as Version of record

PhD Thesis

Cloud Radio Access Network architecture

Towards 5G mobile networks

Aleksandra Checko



Technical University of Denmark
Kgs. Lyngby, Denmark 2016

Cover image: Cloud RUN. Photo by Phil Holmes

Technical University of Denmark
Department of Photonics Enginieering
Networks Technologies and Service Platforms
Ørsteds Plads 343
2800 Kgs. Lyngby
DENMARK
Tel: (+45) 4525 6352
Fax: (+45) 4593 6581
Web: www.fotonik.dtu.dk
E-mail: info@fotonik.dtu.dk

Contents

Abstract

Cloud Radio Access Network (C-RAN) is a novel mobile network architecture which can address a number of challenges that mobile operators face while trying to support ever-growing end-users' needs towards 5th generation of mobile networks (5G). The main idea behind C-RAN is to split the base stations into radio and baseband parts, and pool the Baseband Units (BBUs) from multiple base stations into a centralized and virtualized BBU Pool. This gives a number of benefits in terms of cost and capacity. However, the challenge is then to find an optimal functionality splitting point as well as to design the so-called fronthaul network, interconnecting those parts. This thesis focuses on quantifying those benefits and proposing a flexible and capacity-optimized fronthaul network.

It is shown that a C-RAN with a functional split resulting in a variable bit rate on the fronthaul links brings cost savings due to the multiplexing gains in the BBU pool and the fronthaul network. The cost of a fronthaul network deployment and operation can be further reduced by sharing infrastructure between fronthaul and other services.

The origins of multiplexing gains in terms of traffic burstiness, the tidal effect and various possible functional splits are analyzed and quantified. Sharing baseband resources between many cells is possible for traditional C-RANs. However, in order to further benefit from multiplexing gains on fronthaul, it is recommended to implement a functional split yielding variable bit rate in the fronthaul. For the analyzed data sets, in deployments where diverse traffic types are mixed (bursty, e.g., web browsing and constant bit rate, e.g., video streaming) and cells from various geographical areas (e.g., office and residential) are connected to the BBU pool, the multiplexing gain value reaches six. Using packet-based fronthaul has the potential to utilize fronthaul resources efficiently. However, meeting synchronization and delay requirements is a challenge. As a possible solution, the use of IEEE Precision Time Protocol (PTP) (also known as 1588v2) has been evaluated, and for the analyzed scenario it can assure synchronization on the nanosecond level, fulfilling mobile network requirements. Furthermore, mechanisms to lower delay and jitter have been identified, namely: source scheduling and preemption. An innovative source scheduling scheme which can minimize jitter has been proposed. The scheme is optimized for symmetric downlink and uplink traffic, but can also be used when downlink traffic exceeds uplink. Moreover, a demonstrator of a Software Defined Networking (SDN) controlled Ethernet fronthaul has been built.

Resumé

(Summary in Danish)

Cloud Radio Access Network (C-RAN) er en ny mobilnetarkitektur, som kan imødegå nogle af de udfordringer mobiloperatørerne møder, når de skal understøtte de øgede krav fra brugerne på vej mod 5. generations mobilnet (5G). Idéen bag C-RAN er at opdele basestationerne i radio- og basebandfunktionalitet og så samle basebandfunktionaliteten (de såkaldte *Baseband Units, BBU*) fra flere basestationer i en centraliseret og virtualiseret BBU-pool. Dette giver en række fordele i forhold til kapacitet og omkostninger. Udfordringen er imidlertid så at finde den optimale opdeling af funktionaliteten, samt at designe det såkaldte *fronthaul* net, som forbinder disse dele. Denne afhandling fokuserer på at kvantificere disse fordele og på samme tid foreslå et fleksibelt og kapacitetsoptimeret *fronthaul*-net.

Det vises at et C-RAN med en funktionel opdeling, som giver variabel bitrate på *fronthaul*-forbindelserne, giver omkostningsbesparelser grundet *multiplexing-gain* i både BBU-poolen og i *fronthaul*-nettet. Omkostningerne til udrulning og drift af *fronthaul*-nettet kan reduceres yderligere ved at dele en fælles infrastruktur mellem *fronthaul* og andre tjenester.

Kilderne til *mulitplexing gain* i form af trafik-*burstines*, *tidal effect* og de forskellige måder at opdele funktionaliteten på er blevet analyseret og kvantificeret. Deling af BBU ressourcer mellem mange celler er muligt selv i traditionelle C-RAN. Vil man imidlertid kunne nyde godt af fordelene ved *mulitplexing gain* i *fronthaul*, anbefales det at implementere en funktionel opdeling, som giver variabel bitrate i *fronthaul*-nettet. De data, der er analyseret her, viser et muligt *multiplexing gain* på seks for udrulninger hvor forskellige trafiktyper (bursty, f.eks. web-browsing og konstant bitrate trafik, f.eks. video streaming) og celler fra forskelligartede grografiske områder (f.eks., erhvervs- og beboelsesområder) forbindes til samme BBU-pool. Brug af pakkekoblet fronthaul giver potentielt mulighed for effektiv udnyttelse af fronthaul-ressourcer. Det er imidlertid en udfordring at overholde kravene til synkronisering og forsinkelser. Som en mulig løsning er brug af IEEE Precision Time Protocol (PTP, også kendt som IEEE 1588v2) blevet evalueret og for de analyserede scenarier kan den opfylde synkroniseringskravene. Endvidere er der blevet fundet frem til mekanismer der kan mindske forsinkelse og jitter: *source scheduling* og *preemption*. Der foreslås et nyskabende *source scheduling*-system, som kan minimere *jitter*. Systemet er optimeret til symmetrisk downlink- og uplinktrafik,

men kan også finde anvendelse når downlink overskrider uplink. Slutteligt er der opbygget en demonstrator af et Software Defined Network (SDN) styret, Ethernet baseret fronthaul.

Preface

This dissertation presents a selection of the research work conducted during my PhD study from January 1, 2013 until February 15, 2016 under supervision of Associate Professor Michael Stübert Berger, Associate Professor Henrik Lehrmann Christiansen, Professor Lars Dittmann, Dr Georgios Kardaras (January 2013-December 2013), Dr Bjarne Skak Bossen (January 2014-May 2015), and Dr Morten Høgdal (June 2015-February 2016). It is submitted to the Department of Photonics Engineering at the Technical University of Denmark in a partial fulfillment of the requirements for the Doctor of Philosophy (PhD) degree.

This Industrial PhD project was done in the Networks Technologies and Service Platforms group at the Department of Photonics Engineering at the Technical University of Denmark (DTU), Kgs. Lyngby, Denmark and in MTI Radiocomp, Hillerød, Denmark, where I was employed. The work was co-financed by the Ministry of Higher Education and Science of Denmark within the framework of Industrial PhD Program. The work benefited from collaboration within the EU HARP project, especially from a six-month external stay at Alcatel-Lucent Bell Labs France under the guidance of Mr. Laurent Roullet.

Aleksandra Checko
Kgs. Lyngby, February 2016

x

Acknowledgements

Being a PhD student is not an easy job. Not only does one need to define a topic that is challenging, but also advance the state-of-the-art within it. Whatever one finds, he or she is welcome to re-search. On top of that, various formal aspects of the PhD project need to be taken care of. Fortunately, I was not alone in this battle. I would like to take the opportunity to thank some of the people that made the accomplishment of this thesis possible.

First and foremost, I would like to thank my supervisors: Michael Berger, Bjarne Bossen, Henrik Christiansen, Lars Dittmann, Georgios Kardaras, and Morten Høgdal for their mentoring and support. I am particularly grateful to Michael Berger and to MTI Radiocomp for giving me this opportunity. Special thanks to Henrik Christiansen for making things as simple as possible, but not simpler. Moreover, I am grateful to Lara Scolari and Thomas Nørgaard for their help in starting this project, and to Laurent Roullet, Thomas Nørgaard, and Christian Lanzani for asking interesting questions and helping to shape the directions this thesis took.

I would like to acknowledge the members of the Networks group at DTU Fotonik and colleagues from MTI for all the fruitful discussions, and for maintaining a friendly atmosphere at work. Lunch and seminar discussions were one-of-a-kind.

Merci beaucoup Laurent Roullet for welcoming me in his group at Alcatel Lucent Bell Labs for a six-month external stay, and to Aravinthan Gopalasingham for our fruitful collaboration. I am also grateful to the colleagues from Bell Labs for a warm welcome, guidance and making my time in Paris more enjoyable, especially for our trips to Machu Picchu.

I would like to show my gratitude to Henrik Christiansen, Michael Berger, Morten Høgdal, Andrea Marcano, Matteo Artuso, Jakub Jelonek, Małgorzata Chećko and Georgios Kardaras for reviewing this thesis. Mange tak til Henrik Christiansen for translating the abstract.

This project would be a different experience if it was not for my cooperation with members of the HARP project, the Ijoin project, and the IEEE Student Branch at DTU. I am thankful for this multicultural experience.

And last but not least, I am grateful to my family and friends for their continuous support. A most special thanks to Jakub Jelonek who stood by me in all joyful but also

tough moments.

I had a great opportunity to work with so many bright people, from whom I have learned so much. Dziękuję!

List of Figures

List of Tables

Acronyms

3GPP Third Generation Partnership Project

ACK acknowledgment

AM Acknowledged Mode

ARPU Average Revenue Per User

ARQ Automatic Repeat-reQuest

BB Baseband

BBU Baseband Unit

BLP Bit-level processing

BSE Base Station Emulator

BTS Base Transceiver Station

C-RAN Cloud Radio Access Network

CA Carrier Aggregation

CAPEX CAPital EXpenditure

CDF Cumulative Distribution Function

CDN Content Distribution Network

CMOS Complementary metal-oxide-semiconductor

CoMP Coordinated Multi-Point

CP cyclic-prefix

CPRI Common Public Radio Interface

CPRI2Eth CPRI-to-Ethernet

CRC Cyclic Redundancy Check

DAGC Digital Automatic Gain Control

DCN Dynamic Circuit Network

DL downlink

DPD Digital-to-Analog Converter

DPI Deep Packet Inspection

DSL Digital Subscriber Line

DSN Distributed Service Network

DSP Digital Signal Processor

eICIC enhanced Inter-cell Interference Coordination

EPC Evolved Packet Core

EVM Error Vector Magnitude

FCS Frame check sequence

FDD Frequency-Division Duplex

FEC Forward Error Correction

FFT Fast Fourier Transform

FIFO First Input First Output

FPGA Field-Programmable Gate Array

GOPS Giga Operations Per Second

GPP General Purpose Processor

GPS Global Positioning System

GSM Global System for Mobile Communications

HARQ Hybrid ARQ

HSS Home Subscriber Server

ICI Inter Cell Interference

ICIC Inter-cell Interference Coordination

IFFT Inverse FFT

IoT Internet of Things

IP Internet Protocol

IQ In-phase/Quadrature

KPI Key Performance Indicators

LTE Long Term Evolution

LTE-A LTE-Advanced

M2M Machine to Machine

MAC Media Access Control

MEF Metro Ethernet Forum

MG Multiplexing gain

MIMO Multiple Input Multiple Output

MME Mobility Management Entity

MoU Memorandum of Understanding

MPLS Multiprotocol Label Switching

MPLS-TP MPLS Transport Profile

MTU Maximum Transmission Unit

NACK non-acknowledgement

NFV Network Function Virtualisation

NGFI Next Generation Fronthaul Interface

NGMN Next Generation Mobile Networks

OAM Operations, Administration and Maintenance

OBSAI Open Base Station Architecture Initiative

ODU Optical channel Data Unit

OFDM Orthogonal Frequency-Division Multiple

OPEX OPerating EXpenditure

ORI Open Radio equipment Interface

OTN Optical Transport Network

PA Power Amplifier

PAR Project Authorization Request

PBB-TE Provider Backbone Bridge - Traffic Engineering

PDCP Packet Data Convergence Protocol

PDN-GW Packet Data Network Gateway

PDSCH Physical Downlink Shared Channel

PDU Packet Data Unit

PG Pooling gain

PLL Phase-locked Loop

PON Passive Optical Network

PTP Precision Time Protocol

QAM Quadrature Amplitude Modulation

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

QSNR Quantization SNR

RAN Radio Access Network

RANaaS RAN-as-a-Service

RF Radio Frequency

RLC Radio Link Control

RNC Radio Network Controller

ROHC RObust Header Compression

RRH Remote Radio Head

RRM Radio Resource Management

RTT Round Trip Time

SAE-GW Serving SAE Gateway

SDH Synchronous Digital Hierarchy

SDN Software Defined Networking

SDR Software Defined Radio

SDU Service Data Unit

SLA Service Level Agreement

SNR Signal-to-noise ratio

SoC System on a Chip

SON Self-Organizing Networks

SONET Synchronous Optical NETworking

TCO Total Cost of Ownership

TD-SCDMA Time Division Synchronous Code Division Multiple Access

TDD Time-Division Duplex

TM Transparent Mode

TSN Time-Sensitive Networking

UE User Equipment

UL uplink

UM Unacknowledged Mode

UMTS Universal Mobile Telecommunications System

WDM Wavelength-division multiplexing

WNC Wireless Network Cloud

CHAPTER 1

Introduction

The advancements in Internet connectivity are driving socio-economical changes including personalized broadband services, like TV on demand, laying the ground for e-health, self-driving cars, augmented reality, intelligent houses and cities, connecting various industrial sensors, e.g. for monitoring river levels. Figure 1.1 shows a pertinent example on how common smart mobile phones tuned into and how used people became to living their life with them in their hands.

Figure 1.1: The pope elections in 2005 and 2013.

More and more of our communication becomes mobile. In 1991, the first digital mobile call was made through the Global System for Mobile Communications (GSM) (2G) network by a Finnish prime minister. By 2001, the number of network subscribers exceeded 500 million [1]. That same year, the first 3G network, the Universal Mobile Telecommunications System (UMTS), was introduced to increase the speed of data transmission. The mobile Internet connectivity has gained a wide spread popularity with

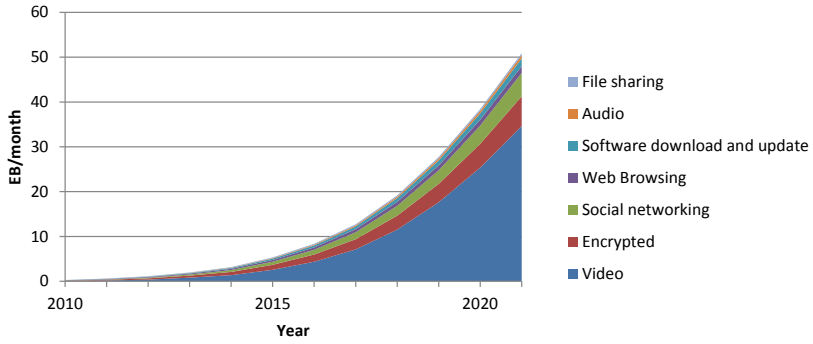


Figure 1.2: Traffic growth in mobile networks. Source: Ericsson, November 2015.

Long Term Evolution (LTE) networks, which have been commercially deployed since 2011. By the end of 2012, the number of LTE subscribers had exceeded 60 million, while the total number of mobile devices amounted to 7 billion [1], exceeding the world's population. The number of subscribers is forecast to grow further, reaching 9 billion in 2021 [2], especially with the increasing popularity of Machine to Machine (M2M) communication.

By 2021, a new generation of mobile networks will be standardized [3] - 5G - to satisfy ever-growing traffic demand, offering increased speed and shorter delays, the latter enabling a tactile Internet. It will be also able to offer ultra-high reliability, and connect a vast number of various devices, such as cars and sensors, to the Internet, forming an Internet of Things (IoT) [4]. The actual applications of the future mobile networks are probably beyond our current imagination.

Historical data as well as traffic forecast shows that this increase in number of subscribers is accompanied by an exponential growth in traffic, with video occupying the most bandwidth [5], [2], as shown on Figure 1.2. In order to support such a growth more cells or higher capacity per cell needs to be provided, which results in increased cost. At the same time, users would like to pay less and use more data. Therefore the increase of cost cannot follow the same trend line as traffic growth. The same applies to power consumption both in terms of cost as well as respect towards the environment. Disruptive, affordable solutions are needed to deliver more capacity, shorter delays, and improved reliability at the same time not increasing the power consumption.

In order to address the capacity requirements more cells can be deployed, or higher capacity needs to be provided per each cell. Small cells can be deployed in places with high user activity. Self-Organizing Networks (SON) techniques are important to ease management and optimization of networks with many cells. In dense deployments interference is a challenge, therefore techniques like enhanced Inter-cell Interference Coordination (eICIC) and Coordinated Multi-Point (CoMP) are essential. To increase the capacity of each cell more antennas can be used using a Massive Multiple Input Multiple Output (MIMO). As spectrum is scarce in currently explored bands, higher frequency bands,

including millimeter waves, are explored. While virtualization became a commodity in IT sector, it is now considered to be applied for mobile networks. Leveraging Network Function Virtualisation (NFV) and SDN offers potential benefits of cost efficient and flexible deployment of both core network and base stations [4]. The latter two techniques are essential enablers for C-RAN. The main concept behind C-RAN is to pool the BBUs from multiple base stations into a centralized and virtualized BBU Pool, while leaving the Remote Radio Heads (RRHs) and antennas at the cell sites. C-RAN enables energy efficient network operation and possible cost savings on baseband resources. Furthermore, it improves network capacity by performing load balancing and cooperative processing of signals originating from several base stations. Therefore base stations with high load and high deployment density may be deployed in C-RAN architecture. For low utilization base stations, e.g. in rural areas, C-RAN may not be the most cost-optimal solution.

Virtualization, now commodity for IT sector, is gaining a momentum in the telecommunication sector. The main drive is cost savings and ability to offer additional services to the users to increase the revenue instead of being just a data carrier. C-RAN exploits virtualization at the air interface, however both mobile core networks and core networks can benefit from cloudification. With a great industry and academia interest in NFV, C-RAN, based on NFV, is an indispensable part of 5G.

1.1 Thesis Outline

C-RAN is a novel RAN architecture that can address a number of challenges that the operators face while trying to support the growing needs of end-users, therefore it is seen as a major technological foundation for 5G networks [[4], [6], [7], [8], [9]]. Chapter 2 provides an extensive introduction to C-RAN, outlining its fundamental aspects, advantages, and the technical solutions that can address the challenges it faces. In addition, future research directions are outlined.

However, for the deployments to be widespread, they need to be economically feasible. Chapter 3 evaluates energy and cost savings coming from the statistical multiplexing gain in C-RAN. Three studies are covered, the first one not including protocol processing, the second exploring a tidal effect, and the third looking at different method of measuring multiplexing gains.

One of the major challenges to address to deploy C-RAN is a flexible fronthaul with optimized capacity requirements. Chapter 4 puts special focus on fronthaul networks, outlining their requirements, evaluating more traditional, circuit-switched transport solution, as well as innovative packet-based transport techniques. Emphasis is placed on Ethernet, packet-based fronthaul, exploring options to ensure clock delivery and to meet delay requirements.

Finally, Chapter 5 concludes the dissertation and explores future research directions.

1.2 A Note on contributions

The work done in this thesis would not be possible without my supervisors and colleagues from EU project HARP consortium. This section highlights contributions of this thesis and describes my role in the joint work.

Summarizing state-of-the-art literature on C-RAN up to beginning of 2014. I have performed an extensive research on the state-of-the-art of C-RAN, sharing my experiences from the time when I was getting familiar with the topic and determining the directions of this thesis. I led writing of the paper [10], wrote myself all the sections apart from section III.C and VIII as well as revised the whole paper including the input from co-authors. Given the statistics on downloads (3500+) and citations (30+) within a year from publication, I trust that the community found it useful.

Evaluating C-RAN benefits in terms of exploring statistical multiplexing gain. I have designed and performed OPNET simulations on evaluating statistical multiplexing gains in [11], [12] and [13], based on the measurement method proposed by me before commercing on the Ph.D. project [14]. The work presented in [13] was done with Andrijana P. Avramova, who performed teletraffic studies. All the papers were prepared under the supervision of Henrik Christiansen and Michael Berger.

Summarizing requirements and challenges of Ethernet-based fronthaul. I have organized requirements and analyzed factors that are challenging for achieving synchronization in packet-based C-RAN fronthaul. Via discrete event-based simulations I performed a feasibility study showing the performance for frequency and phase synchronization using IEEE 1588 in Ethernet networks in [15]. The work was done under the supervision of Henrik Christiansen and Michael Berger.

Building Ethernet - based, SDN - controlled fronthaul ready for a single CPRI stream. This work was done in Alcatel-Lucent Bell Labs France in collaboration with Aravinthan Gopalasingham. We built this demonstrator configuring Open Flow switches and OpenDaylight together, while I was responsible for configuring DPDK traffic generation and performing delay measurements.

Proposing source scheduling algorithm. I proposed a traffic scheduling algorithm, important for enabling sharing Ethernet network between many fronthaul streams.

Experimentally verifying feasibility of using OTN for C-RAN. Together with Georgios Kardaras and Altera we performed lab measurements on feasibility of applying OTN to CPRI/OBSAI fronthaul. MTI Radiocomp setup was used on CPRI/OBSAI layer, while Altera provided OTN equipment.

1.3 Publications prepared during the course of the PhD project

1.3.1 Journals

1. A. Checko[1st], A.P. Avramova[1st], M.S. Berger and H.L. Christiansen, "Evaluating C-RAN fronthaul functional splits in terms of network level energy and cost savings",

accepted to IEEE Journal Of Communications And Networks

2. A. Checko, H. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M.S. Berger and L. Dittmann, “Cloud RAN for mobile networks - a technology overview” In: IEEE Surveys and Tutorials journal, Vol. 17, no. 1, (Firstquarter 2015), pp. 405–426

3. M. Artuso, D. Boviz, A. Checko, et al. , “Enhancing LTE with Cloud-RAN and Load-Controlled Parasitic Antenna Arrays”, submitted to IEEE Communications Magazine

1.3.2 Conferences

1. A. Checko, A. Juul, H. Christiansen and M.S. Berger, "Synchronization Challenges in Packet-based Cloud-RAN Fronthaul for Mobile Networks", In proceedings of IEEE ICC 2015 - Workshop on Cloud-Processing in Heterogeneous Mobile Communication Networks (IWCPM)

2. H. Holm, A. Checko, R. Al-obaidi and H. Christiansen, “Optimizing CAPEX of Cloud-RAN Deployments in Real-life Scenarios by Optimal Assignment of Cells to BBU Pools”, in Networks and Communications (EuCNC), 2015 European Conference on, pp.205-209, June 29 2015-July 2 2015

3. R. Al-obaidi, A. Checko, H. Holm and H. Christiansen, “Optimizing Cloud-RAN Deployments in Real-life Scenarios Using Microwave Radio”, in Networks and Communications (EuCNC), 2015 European Conference on , vol., no., pp.159-163, June 29 2015-July 2 2015

4. L. Dittmann, H.L. Christiansen and A. Checko, “Meeting fronthaul challenges of future mobile network deployments - the HARP approach”, invited to WONC Workshop, In proceedings of IEEE GLOBECOM 2014, Austin/USA, 2 December 2014

5. A. Checko, H. Holm, and H. Christiansen, “Optimizing small cell deployment by the use of C-RANs,” in European Wireless 2014; 20th European Wireless Conference; Proceedings of, pp.1-6, 14-16 May 2014

6. S. Michail, A. Checko and L. Dittmann, “Traffic Adaptive Base Station Management Scheme for Energy-Aware Mobile Networks”, poster presented at EuCNC 2014

7. A. Checko, H.L. Christiansen and M.S. Berger, "Evaluation of energy and cost savings in mobile Cloud Radio Access Networks", In proceedings of OPNETWORK 2013 conference, Washington D.C., USA, August 26-30 2013

8. A. Dogadaev, A. Checko, A.P. Avramova, A. Zakrzewska, Y. Yan, S. Ruepp, M.S. Berger, and L. Dittmann and H. Christiansen, “Traffic Steering Framework for Mobile-Assisted Resource Management in Heterogeneous Networks”, In proceedings of IEEE 9th International Conference on Wireless and Mobile Communications (ICWMC) 2013, Nice, France 21-26 July 2013.

1.3.3 Patent application

1. A. Checko "Method for Radio Resource Scheduling”. US patent provisional application number 62294957, filled on 12.02.2016

1.3.4 Others

1. A. Checko, G. Kardaras, C. Lanzani, D. Temple, C. Mathiasen, L. Pedersen, B. Klaps, “OTN Transport of Baseband Radio Serial Protocols in C-RAN Architecture for Mobile Network Applications”. Industry white paper, March 2014

2. Contribution to the deliverable of the EU project HARP D6.4 “Eth2CPRI prototype implementation”, presented to European Commission in November 2015

3. Contribution to the deliverable of the EU project HARP D6.3 “Protocol Extensions Design and Implementation”, presented to European Commission in August 2015

4. Coordination and contribution to Deliverable 6.2 “Aggregation network optimization” of the EU project HARP, presented to European Commission in January 2015

5. Contribution to Y2 report of the EU project HARP, presented to European Commission in November 2014

6. Contribution to Y1 report of the EU project HARP, presented to European Commission in November 2013

7. Contribution to the deliverable of the EU project HARP D6.1 “Aggregation Network Definition”, presented to European Commission in May 2014

8. Contribution to the deliverable of the EU project HARP D3.1 Requirements, metrics and network definition, presented to European Commission in July 2013

9. Contribution to the deliverable of the EU project HARP D1.5 “Final plan for the use and dissemination of Foreground”, presented to European Commission partners in May 2014

CHAPTER 2

C-RAN overview

Base Transceiver Station (BTS), NodeB, eNodeB. Those are the names used to describe a base station in GSM, UMTS and LTE standards, respectively. As a concept, and logical node, a base station is responsible for receiving signal from/sending to user, prepare it to be send up to/received from the core network and organize transmission. Physically, this node can be deployed as a standalone base station, base station with RRH or Cloud RAN (C-RAN).

As spectral efficiency for the LTE standard is approaching the Shannon limit, the most prominent way to increase network capacity is by either adding more cells, creating a complex structure of Heterogeneous and Small cell Networks (HetSNets) [16] or by implementing techniques such as multiuser MIMO [17] as well as Massive MIMO [18], where numerous antennas simultaneously serve a number of users in the same time-frequency resource. However, this results in growing inter-cell interference levels and high costs.

The Total Cost of Ownership (TCO) in mobile networks includes CAPital EXpenditure (CAPEX) and OPerating EXpenditure (OPEX). CAPEX mainly refers to expenditure relevant to network construction which may span from network planning to site acquisition, RF hardware, baseband hardware, software licenses, leased line connections, installation, civil cost and site support, like power and cooling. OPEX covers the cost needed to operate the network, i.e. site rental, leased line, electricity, operation and maintenance as well as upgrade [19]. CAPEX and OPEX are increasing significantly when more base stations are deployed. More specifically, CAPEX increases as base stations are the most expensive components of a wireless network infrastructure, while OPEX increases as, among others, cell sites demand a considerable amount of power to operate, e.g., China Mobile estimates 72% of total power consumption originates from the cell sites [20]. Mobile network operators need to cover the expenses for network construction, operation, maintenance and upgrade; meanwhile, the Average Revenue Per User (ARPU) stays flat or even decreases over time, as the typical user is more and more data-hungry but expects to pay less for data usage.

C-RAN is a novel mobile network architecture, which has the potential to answer the above mentioned challenges. The concept was first proposed in [21] and described in detail in [20]. In C-RAN, baseband processing is centralized and shared among sites in a virtualized BBU Pool. This means that it is well prepared to adapt to non-uniform traffic and utilizes resources, i.e., base stations, more efficiently. Due to that fact that fewer BBUs are needed in C-RAN compared to the traditional architecture, C-RAN also

has the potential to decrease the cost of network operation, because power and energy consumption are reduced compared to the traditional RAN architecture. New BBUs can be added and upgraded easily, thereby improving scalability and easing network maintenance. A virtualized BBU Pool can be shared by different network operators, allowing them to rent Radio Access Network (RAN) as a cloud service. As BBUs from many sites are co-located in one pool, they can interact with lower delays – therefore mechanisms introduced for LTE-Advanced (LTE-A) to increase spectral efficiency and throughput, such as eICIC and CoMP are greatly facilitated. Methods for implementing load balancing between cells are also facilitated. Furthermore, network performance is improved, e.g., by reducing delay during intra-BBU Pool handover.

C-RAN architecture is targeted by mobile network operators, as envisioned by China Mobile Research Institute [20], IBM [21], Alcatel-Lucent [22], Huawei [23], ZTE [24], Nokia Siemens Networks [19], Intel [25] and Texas Instruments [26]. Moreover, C-RAN is seen as a typical realization of mobile network supporting soft and green technologies in fifth generation (5G) mobile networks [27]. However, C-RAN is not the only candidate architecture that can answer the challenges faced by mobile network operators. Other solutions include small cells, being part of HetSNets and Massive MIMO. Small cells deployments are the main competitors for outdoor hot spot as well as indoor coverage scenarios. All-in-one small footprint solutions like Alcatel-Lucent's LightRadio [28] can host all base station functionalities in a several-liter box. They can be placed outdoors reducing cost of operation associated to cooling and cell site rental. However, they will be underutilized during low-activity periods and cannot employ collaborative functionalities as well as C-RAN. Moreover, they are more difficult to upgrade and repair than C-RAN. A brief comparison between C-RAN, Massive MIMO and HetSNets is outlined in [16]. Liu *et al.* in [29] prove that energy efficiency of large scale Small Cell Networks is higher compared with Massive MIMO. Furthermore, cost evaluation on different options needs to be performed in order for a mobile network operator to choose an optimal solution. Comparison of TCO including CAPEX and OPEX over 8 years of a traditional LTE macro base station, LTE C-RAN and LTE small cell shows that the total transport cost per Mbps is highest for macro cell deployments - 2200$, medium for C-RAN - 1800$ and 3 times smaller for small cell - 600$ [30]. Therefore the author concludes that C-RAN needs to achieve significant benefits to overcome such a high transportation cost. Collaborative techniques such as CoMP and eICIC can be implemented in small cells giving higher benefits in HetNet configuration instead of C-RAN. The author envisions that C-RAN might be considered for special cases like stadium coverage. However, C-RAN is attractive for operators that have free/cheap fiber resources available.

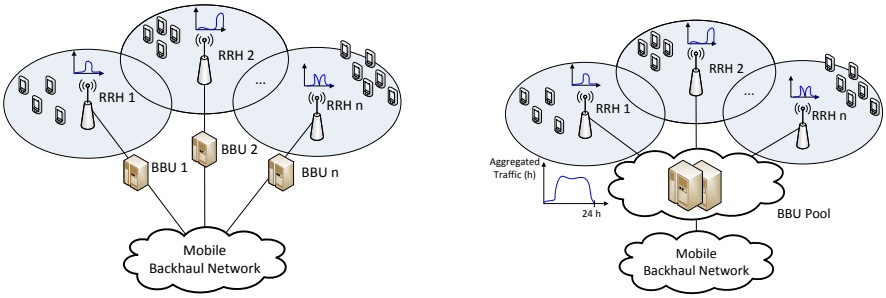
This chapter surveys the state-of-the-art literature published on C-RAN and its implementation until 2014. Such input helps mobile network operators to make an optimal choice on deployment strategies. The chapter is organized as follows. In Section 2.1 the fundamental aspects of C-RAN architecture are introduced. Moreover, in Section 2.2 the advantages of this architecture are discussed in detail along with the challenges that need to be overcome before fully exploiting its benefits in Section 2.3. In Section 2.4 a

number of constraints in regards to the transport network capacity imposed by C-RAN are presented and possible solutions are discussed, such as the utilization of compression schemes. In Sections 2.5, 2.6 an overview of the state-of-the-art hardware solutions that are needed to deliver C-RAN from the radio, baseband and network sides are given. As the BBU Pool needs to be treated as a single entity, in Section 2.7 an overview of virtualization techniques that can be deployed inside a BBU Pool are presented. In Section 2.8 possible deployment scenarios of C-RAN are evaluated. In Section 2.9 ongoing work on C-RAN and examples of first field trials and prototypes are summarized. Furthermore, future research directions are outlined in Section 2.10. Section 2.11 summarizes the chapter.

This section, Sections 2.1-2.6 and 2.8-2.9 were originally published in [10]. This chapter provides an updated and extended version to reflect recent developments, especially in case of Virtualization, covered in Section 2.7 and showing latest trends in 2.2.5 and 2.9.4.

2.1 What is C-RAN? Base Station architecture evolution

C-RAN is a network architecture where baseband resources are pooled, so that they can be shared between base stations. Figure 2.1 gives an overview of the overall C-RAN architecture. This section gives an introduction to base station evolution and the basis of the C-RAN concept.



(a) RAN with RRH (b) C-RAN

Figure 2.1: Statistical multiplexing gain in C-RAN architecture for mobile networks.

The area which a mobile network covers is divided into cells, therefore mobile networks are often called cellular networks. Traditionally, in cellular networks, users communicate with a base station that serves the cell under coverage of which they are

located. The main functions of a base station can be divided into baseband processing and radio functionalities. The main sub-functions of baseband processing module are shown in left side of Figure 2.2. Among those one can find coding, modulation, Fast Fourier Transform (FFT), etc. The radio module is responsible for digital processing, frequency filtering and power amplification.

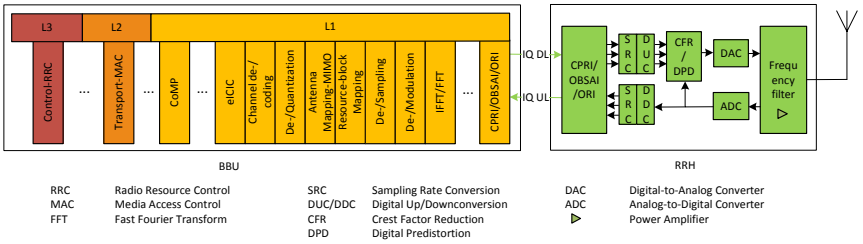


Figure 2.2: Base station functionalities. Exemplary baseband processing functionalities inside BBU are presented for LTE implementation. Connection to RF part and sub modules of RRH are shown.

2.1.1 Traditional architecture

In the traditional architecture, radio and baseband processing functionality is integrated inside a base station. The antenna module is generally located in the proximity (few meters) of the radio module as shown in Figure 2.3a as coaxial cables employed to connect them exhibit high losses. This architecture was popular for 1G and 2G mobile networks deployment.

2.1.2 Base station with RRH

In a base station with RRH architecture, the base station is separated into a radio unit and a signal processing unit, as shown in Figure 2.3b. The radio unit is called a RRH or Remote Radio Unit (RRU). RRH provides the interface to the fiber and performs digital processing, digital to analog conversion, analog to digital conversion, power amplification and filtering [31]. The baseband signal processing part is called a BBU or Data Unit (DU). More about BBU can be found in Chapter 16 of [32]. Interconnection and function split between BBU and RRH are depicted in Figure 2.2. This architecture was introduced when 3G networks were being deployed and right now the majority of base stations use it.

The distance between a RRH and a BBU can be extended up to 40 km, where the limitation is coming from processing and propagation delay as explained in Section 4.1.3. Optical fiber and microwave connections can be used. In this architecture, the BBU equipment can be placed in a more convenient, easily accessible place, enabling cost savings on site rental and maintenance compared to the traditional RAN architecture,

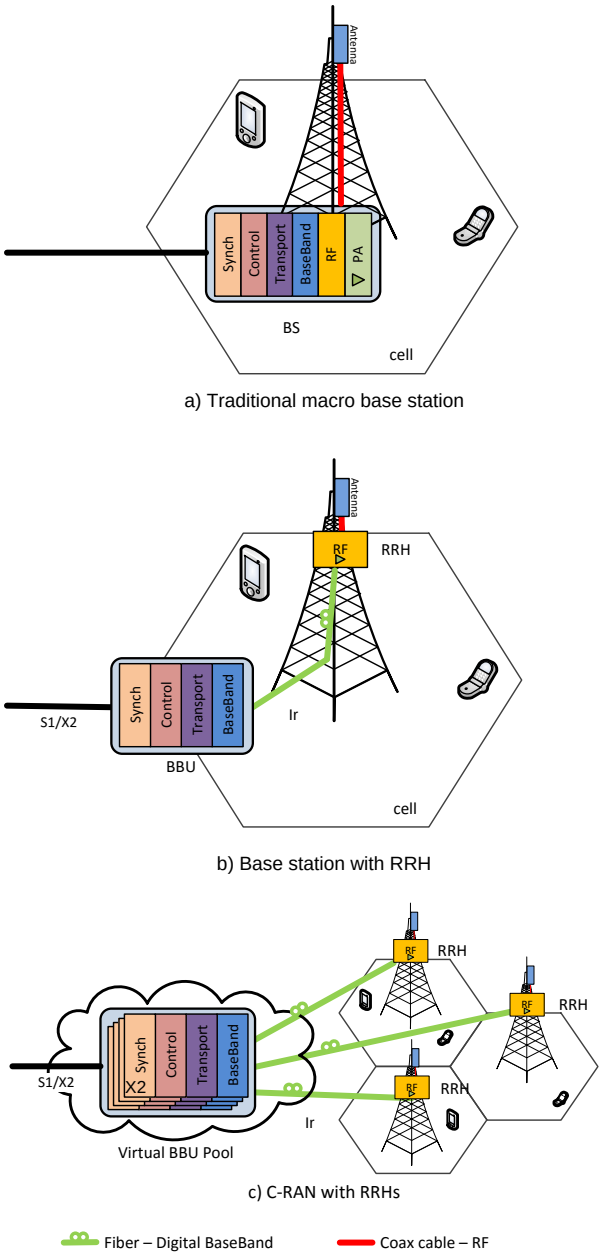


Figure 2.3: Base station architecture evolution.

where a BBU needs to be placed close to the antenna. RRHs can be placed up on poles or rooftops, leveraging efficient cooling and saving on air-conditioning in BBU housing. RRHs are statically assigned to BBUs similarly to the traditional RAN. One BBU can serve many RRHs. RRHs can be connected to each other in a so called daisy-chained architecture. An Ir interface is defined, which connects RRH and BBU.

Common Public Radio Interface (CPRI) [33] is the radio interface protocol widely used for In-phase/Quadrature (IQ) data transmission between RRHs and BBUs - on Ir interface. It is a constant bit rate, bidirectional protocol that requires accurate synchronization and strict latency control. Other protocols that can be used are Open Base Station Architecture Initiative (OBSAI) [34] and Open Radio equipment Interface (ORI) [35], [36]. For LTE base stations the X2 interface is defined between base stations, the S1 interface connects a base station with mobile core network.

2.1.3 Cloud base station architecture - C-RAN

In C-RAN, in order to optimize BBU utilization between heavily and lightly loaded base stations, the BBUs are centralized into one entity that is called a BBU/DU Pool/Hotel. A BBU Pool is shared between cell sites and virtualized as shown in Figure 2.3c. A BBU Pool is a virtualized cluster which can consist of general purpose processors to perform baseband (PHY/MAC) processing. X2 interface in a new form, often referred to as X2+ organizes inter-cluster communication as well as communication to other pools.

The concept of C-RAN was first introduced by IBM [21] under the name Wireless Network Cloud (WNC) and builds on the concept of Distributed Wireless Communication System [37]. In [37] Zhou *et al.* propose a mobile network architecture in which a user communicates with densely placed distributed antennas and the signal is processed by Distributed Processing Centers (DPCs). C-RAN is the term used now to describe this architecture, where the letter C can be interpreted as: Cloud, Centralized processing, Cooperative radio, Collaborative or Clean.

Figure 2.4 shows an example of a C-RAN mobile LTE network. The fronthaul part of the network spans from the RRHs sites to the BBU Pool. The backhaul connects the BBU Pool with the mobile core network. At a remote site, RRHs are co-located with the antennas. RRHs are connected to the high performance processors in the BBU Pool through low latency, high bandwidth optical transport links. Digital baseband, i.e., IQ samples, are sent between a RRH and a BBU.

Table 2.1 compares traditional base station, base station with RRH and base station in C-RAN architecture.

2.2 Advantages of C-RAN

Both macro and small cells can benefit from C-RAN architecture. For macro base station deployments, a centralized BBU Pool enables an efficient utilization of BBUs and reduces the cost of base stations deployment and operation. It also reduces power consumption and

Table 2.1: Comparison between traditional base station, base station with RRH and C-RAN

Architecture	Radio and baseband functionalities	Problem it addresses	Problems it causes
Traditional base station	Co-located in one unit	-	High power consumption Resources are underutilized
Base station with RRH	Spitted between RRH and BBU. RRH is placed together with antenna at the remote site. BBU located within 20-40 km away. Generally deployed nowadays	Lower power consumption. More convenient placement of BBU	Resources are underutilized
C-RAN	Spitted into RRH and BBU. RRH is placed together with antenna at the remote site. BBUs from many sites are co-located in the pool within 20-40 km away. Field trials and early deployments (2015)	Even lower power consumption. Lower number of BBUs needed - cost reduction	Considerable transport resources between RRH and BBU

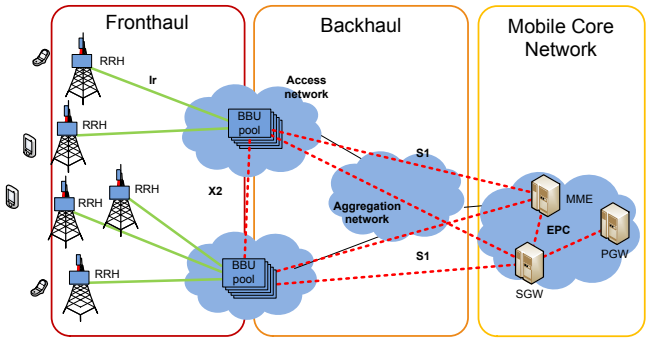


Figure 2.4: C-RAN LTE mobile network.

provides increased flexibility in network upgrades and adaptability to non-uniform traffic. Furthermore, advanced features of LTE-A, such as CoMP and interference mitigation, can be efficiently supported by C-RAN, which is essential especially for small cells deployments. Last but not least, having high computational processing power shared by many users placed closer to them, mobile operators can offer users more attractive Service Level Agreements (SLAs), as the response time of application servers is noticeably shorter if data is cached in BBU Pool [38]. Network operators can partner with third-party service developers to host servers for applications, locating them in the cloud - in the BBU Pool [39]. In this section advantages of C-RAN are described and motivated: A. Adaptability to nonuniform traffic and scalability, B. Energy and cost savings, C. Increase of throughput, decrease of delays as well as D. Ease in network upgrades and maintenance.

2.2.1 Adaptability to nonuniform traffic and scalability

Typically, during a day, users are moving between different areas, e.g., residential and office. Figure 2.5 illustrates how the network load varies throughout the day. Base stations are often dimensioned for busy hours, which means that when users move from office to residential areas, the huge amount of processing power is wasted in the areas from which the users have moved. Peak traffic load can be even 10 times higher than during off-the-peak hours [20]. In each cell, daily traffic distribution varies, and the peaks of traffic occur at different hours. Since in C-RAN baseband processing of multiple cells is carried out in the centralized BBU pool, the overall utilization rate can be improved. The required baseband processing capacity of the pool is expected to be smaller than the sum of capacities of single base stations. The ratio of sum of single base stations capacity to the capacity required in the pool is called statistical multiplexing gain.

In [40] an analysis on statistical multiplexing gain is performed as a function of cell layout. The analysis shows that in the Tokyo metropolitan area, the number of BBUs can be reduced by 75% compared to the traditional RAN architecture. In [41] Werthmann *et al.* prove that the data traffic influences the variance of the compute resource utilization, which in consequence leads to significant multiplexing gains if multiple sectors are

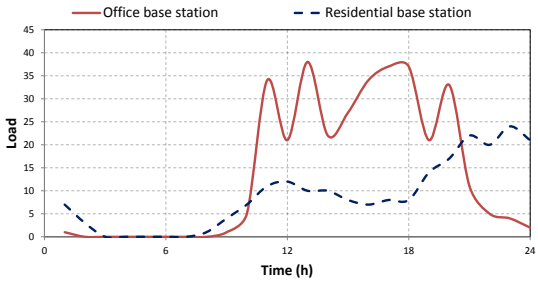


Figure 2.5: Daily load on base stations varies depending on base station location. Data source: [20].

aggregated into one single cloud base station. Aggregation of 57 sectors in a single BBU Pool saves more than 25% of the compute resources. Moreover, the user distribution has a strong influence on the utilization of the compute resources. The results of last three works converge giving around 25% of potential savings on baseband resources. In [42] Bhaumik *et al.* show that the centralized architecture can potentially result in savings of at least 22% in compute resources by exploiting the variations in the processing load across base stations. In [40] Namba *et al.* analyze statistical multiplexing gain as a function of cell layout. The analysis shows that for the metropolitan area, the number of BBUs can be reduced by 75% compared to the traditional base station. In [43] Madhavan *et al.* quantify the multiplexing gain of consolidating WiMAX base stations in different traffic conditions. The gain increases linearly with network size and it is higher when base stations are experiencing higher traffic intensity. On the contrary, in [44] Liu *et al.* analyzed that lighter load can increase the statistical multiplexing gain in virtual base station pool. Moreover, multiplexing gain reaches significant level even for the medium-size pools and the increase in gain for larger pools in negligible.

In author's previous work [11] initial evaluation of statistical multiplexing gain of BBUs in C-RAN was presented. The paper concludes that 4 times less BBUs are needed for user data processing in a C-RAN compared to a traditional RAN for specific traffic patterns, making assumptions of the number of base stations serving different types of areas. The model does not include mobile standard protocols processing. After including protocol processing in [12] the statistical multiplexing gain varied between 1.2 and 1.6 depending on traffic mix, reaching the peak for 30% of office and thereby 70% of residential base stations, thereby enabling saving of 17% - 38%. Those results are presented in Sections 3.2 and 3.3, respectively. Secondly, results obtained via simulations [12] have been compared to the ones achieved with teletraffic theory [45].

All those works referred to the traditional - Baseband (BB)/RF - functional split of

C-RAN. In [46] authors estimate what they define as statistical multiplexing convergence ratio on fronthaul links by averaging observed daily throughput. Calculated ratio equals to three. However, the analysis took only average network load into account and therefore can be interpreted mostly as an average daily fronthaul utilization. In author's most recent work [13] different functional splits (as presented in Section 2.10.1) and different, precisely defined application mixes are investigated. A numerical evaluation was given supporting the intuitive conclusions that the maximum multiplexing gain on BBU resources can be achieved for a fully centralized C-RAN. The more functionality is moved from the BBU pool to the cell site, the lower the multiplexing gain on the BBU pool. However, when traffic starts to be variable bit rate, a multiplexing gain on fronthaul links can be achieved, lowering the required capacity. Those results are presented in Section 3.4.

Statistical multiplexing gain can be maximized by employing a flexible, reconfigurable mapping between RRH and BBU adjusting to different traffic profiles [47]. Statistical multiplexing gain depends on the traffic, therefore it can be maximized by connecting RRHs with particular traffic profiles to different BBU Pools [12]. More on multiplexing gain evaluation can be found in Chapter 3.

Coverage upgrades simply require the connection of new RRHs to the already existing BBU Pool. To enhance network capacity, existing cells can then be split, or additional RRHs can be added to the BBU Pool, which increases network flexibility. Deployment of new cells is in general more easily accepted by local communities, as only a small device needs to be installed on site (RRH) and not a bulky base station. If the overall network capacity shall be increased, this can be easily achieved by upgrading the BBU Pool, either by adding more hardware or exchanging existing BBUs with more powerful ones.

As BBUs from a large area will be co-located in the same BBU Pool, load balancing features can be enabled with advanced algorithms on both the BBU side and the cells side. On the BBU side, BBUs already form one entity, therefore load balancing is a matter of assigning proper BBU resources within a pool. On the cells side, users can be switched between cells without constraints if the BBU Pool has capacity to support them, as capacity can be assigned dynamically from the pool.

2.2.2 Energy and cost savings coming from statistical multiplexing gain in BBU Pool and use of virtualization

By deploying C-RAN, energy, and as a consequence, cost savings, can be achieved [48]. 80% of the CAPEX is spent on RAN [20], therefore it is important to work towards reducing it.

Energy in mobile network is spent on power amplifiers, supplying RRH and BBU with power and air conditioning. 41% of OPEX on a cell site is spent on electricity [20]. Employing C-RAN offers potential reduction of electricity cost, as the number of BBUs in a C-RAN is reduced compared to a traditional RAN. Moreover, in the lower traffic period, e.g., during the night, some BBUs in the pool can be switched off not affecting overall network coverage. Another important factor is the decrease of cooling resources, which

takes 46% of cell site power consumption [20]. Due to the usage of RRHs air conditioning of radio module can be decreased as RRHs are naturally cooled by air hanging on masts or building walls, as depicted in Figure 2.3. ZTE estimates that C-RAN enables 67%-80% power savings compared with traditional RAN architecture, depending on how many cells one BBU Pool covers [24], which stays in line with China Mobile research claiming 71% power savings [49].

Civil work on remote sites can be reduced by gathering equipment in a central room, what contributes to additional OPEX savings.

In total, 15% CAPEX and 50% OPEX savings are envisioned comparing to RAN with RRH [49] or traditional RAN architecture [50]. However, the cost of leasing the fiber connection to the site may increase CAPEX. IQ signal transported between RRHs and BBUs brings up a significant overhead. Consequently, the installation and operation of transport network causes considerable costs for operators.

Moreover, virtualization helps to reduce cost of network deployment and operation, at the same time enabling operators to offer additional services, not only serve as pipelines for carrying user date.

2.2.3 Increase of throughput, decrease of delays

eICIC [51] and CoMP [52] are important features of LTE-A that aim at minimizing inter cell interference and utilizing interference paths constructively, respectively.

If all the cells within a CoMP set are served by one BBU Pool, then a single entity doing signal processing enables tighter interaction between base stations. Therefore interference can be kept to lower level and consequently the throughput can be increased [48]. It has been proven that combining clustering of cells with CoMP makes more efficient use of the radio bandwidth [53]. Moreover, Inter-cell Interference Coordination (ICIC) can be implemented over a central unit - BBU Pool - optimizing transmission from many cells to multiple BBUs [54].

In [55] Huiyu *et al.* discuss the factors affecting the performance of CoMP with LTE-A in C-RAN uplink (UL), i.e., receiver algorithm, reference signals orthogonality and channel estimation, density and size of the network. In [20] authors present simulation results which compare spectrum efficiency of intra-cell and inter-cell JT to non-cooperative transmission. 13% and 20% increase in spectrum efficiency was observed, respectively. For a cell edge user, spectrum efficiency can increase by 75% and 119%, respectively. In [56] Li *et al.* introduce LTE UL CoMP joint processing and verify its operation on a C-RAN test bed around Ericsson offices in Beijing Significant gain was achieved at the cell edge both for intra-site CoMP and inter-site CoMP. Throughput gain is 30-50% when there is no interference and can reach 150% when interference is present. The authors have compared MRC (Maximum Ratio Combining) and full IRC (Interference Rejection Combining). Due to the reduction of X2 usage in C-RAN, real time CoMP can give 10-15% of joint processing gain, while real time ICIC enables 10-30% of multi cell Radio Resource Management (RRM) gain [19]. Performance of multiple-point JT

and multiple-user joint scheduling has been analyzed for a non-ideal channel with carrier frequency offset [57]. When carrier frequency offset does not exceed $\pm 3 \sim 5ppb$, C-RAN can achieve remarkable performance gain on both capacity and coverage even in non-ideal channel, i.e., 20%/52% for cell average/cell edge.

With the introduction of the BBU Pool cooperative techniques, as Multi-Cell MIMO [58] can be enhanced. This can be achieved due to tighter cooperation between base station within a pool. In [59], Liu *et al.* present a downlink Antenna Selection Optimization scheme for MIMO based on C-RAN that showed advantages over traditional antenna selection schemes.

2.2.3.1 Decrease of the delays

The time needed to perform handovers is reduced as it can be done inside the BBU Pool instead of between eNBs. In [60] Liu *et al.* evaluate the improvement on handover performance in C-RAN and compare it with RAN with RRHs. In GSM, the total average handover interrupt time is lower and the signaling is reduced due to better synchronization of BBUs. In UMTS signaling, Iub transport bearer setup and transport bandwidth requirements are reduced, however, the performance improvement may not be sensed by the user. For LTE X2-based inter-eNB handover the delay and failure rate are decreased. Moreover, the general amount of signaling information sent to core mobile network is reduced, after being aggregated in the pool.

2.2.4 Ease in network upgrades and maintenance

C-RAN architecture with several co-located BBUs eases network maintenance: not only C-RAN capacity peaks and failure might be absorbed by BBU Pool automatic reconfiguration, therefore limiting the need for human intervention, but whenever hardware failures and upgrades are really required, human intervention is to be done only in a very few BBU pool locations. On the contrary for traditional RAN, the servicing may be required at as many cell sites as there are in the network. C-RAN with a virtualized BBU Pool gives a smooth way for introducing new standards, as hardware needs to be placed in few centralized locations. Therefore deploying it can be considered by operators as a part of their migration strategy.

Co-locating BBUs in BBU Pool enables more frequent CPU updates than in case when BBUs are located in remote sites. It is therefore possible to benefit from the IT technology improvements in CPU technology, be it frequency clock (Moore's law) or energy efficiency (as currently seen in Intel mobile processor road map or ARM architecture).

Software Defined Radio (SDR) is a well known technology that facilitates implementation in software of such radio functions like modulation/demodulation, signal generation, coding and link-layer protocols. The radio system can be designed to support multiple standards [61]. A possible framework for implementing software base stations that are remotely programmable, upgradable and optimizable is presented in [62]. With such

technology, C-RAN BBU Pool can support multi-standard multi-system radio communications configured in software. Upgrades to new frequencies and new standards can be done through software updates rather than hardware upgrades as it is often done today on non-compatible vertical solutions. Multi-mode base station is therefore expected to alleviate the cost of network development and Operations, Administration and Maintenance (OAM).

2.2.5 Benefits driving deployments

A recent (July 2015) study conducted by LightReading [63] on "What is the most important driver for operators to deploy a centralized or cloud RAN" shows that 27% of respondents choose OPEX reduction through centralization, while another 24% pointed out CAPEX reduction through NFV. The complete results (excluding 3% of responses stating "other") are presented in Figure 2.6. What is interesting is that all the advantages mentioned in previous sections are mentioned apart from boosting cooperative techniques, at least they are not mentioned explicitly. OPEX reduction through centralization was a motivation to study multiplexing gain, as presented in Chapter 3, therefore the result of the survey provers that it was an important parameter to study.

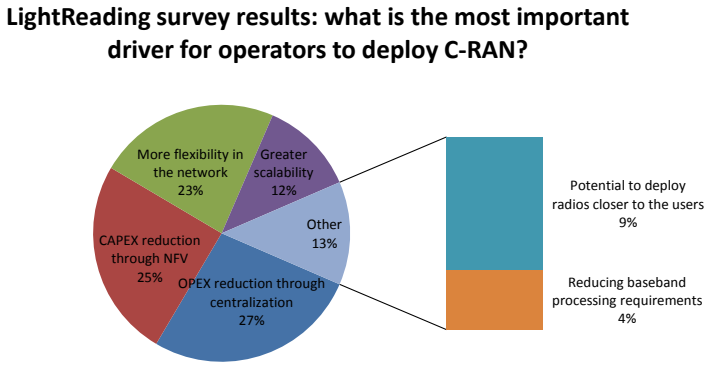


Figure 2.6: Results of the survey [63] on operators' drives for deploying C-RAN.

2.3 Challenges of C-RAN

Before the commercial deployment of C-RAN architectures a number of challenges need to be addressed: A. High bandwidth, strict latency and jitter as well as low cost transport

network needs to be available, B. Techniques on BBU cooperation, interconnection and clustering need to be developed as well as C. Virtualization techniques for BBU Pool need to be proposed. This section elaborates on those challenges. The latter sections present an ongoing work on possible technical solutions that enable C-RAN implementation (Section 2.4, 2.5, 2.6 and 2.7). Figure 2.7 gives an overview of technical solutions addressed in this chapter.

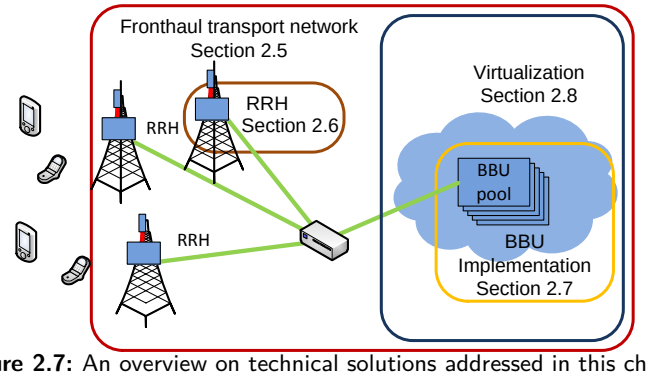


Figure 2.7: An overview on technical solutions addressed in this chapter.

2.3.1 A need for high bandwidth, strict latency and jitter as well as low cost transport network

The C-RAN architecture brings a huge overhead on the optical links between RRH and BBU Pool. Comparing with backhaul requirements, the one on fronthaul are envisioned to be 50 times higher [54].

IQ data is sent between BBU and RRH as shown in Figure 2.2. The main contributors to the size of IQ data are: turbocoding (e.g., in UMTS and LTE 1:3 turbocode is used resulting in three times overhead), chosen radio interface (e.g., CPRI) IQ sample width and oversampling of LTE signal. For example, 30.72 MHz sampling frequency is standardized for 20 MHz LTE, which is more than 20 MHz needed according to Nyquist - Shannon sampling theorem. Total bandwidth depends also on number of sectors and MIMO configuration. Equation 2.1 summarizes factors that influence IQ bandwidth. Scenario of 20 MHz LTE, 15+1 CPRI IQ Sample width, 10/8 line coding, 2x2 MIMO transmission resulting in 2.4576 Gbps bit rate in fronthal link is often treated as a baseline scenario. Consequently, for 20 MHz 4x4 MIMO, 3 sector base station, the expected IQ throughput exceeds 10 Gbps. Examples on expected IQ bit rate between cell site and BBU in LTE-A, LTE, Time Division Synchronous Code Division Multiple Access (TD-SCDMA) and GSM networks can be found in Table 2.2. The centralized BBU Pool should support 10 - 1000 base station sites [20], therefore a vast amount of data needs to be carried towards it.

$$IQBandwidth = samplingFrequency \cdot sampleWidth \cdot 2 \cdot lineCoding \cdot MIMO \cdot noOfSectors \quad (2.1)$$

Table 2.2: IQ bit rates between a cell site and centralized BBU Pool

Cell configuration	Bit rate	Source
20 MHz LTE, 15+1 CPRI IQ Sample width, 10/8 line coding, 2x2 MIMO	2.5 Gbps	
5x20 MHz LTE-A, 15 CPRI IQ Sample width, 2x2 MIMO, 3 sectors	13.8 Gbps	[64]
20 MHz LTE, 4x2 MIMO, 3 sectors	16.6 Gbps	[22]
TD-LTE, 3 sectors	30 Gbps	[65]
1.6 MHz TD-SCDMA, 8Tx/8Rx antennas, 4 times sampling rate	330 Mbps	[20]
TD-SCDMA S444, 3 sectors	6 Gbps	[65]
200 kHz GSM, 2Tx/2Rx antennas, 4x sampling rate	25.6 Mbps	[20]

The transport network not only needs to support high bandwidth and be cost efficient, but also needs to support strict latency and jitter requirements. Below different constraints on delay and jitter are summarized:

1. The most advanced CoMP scheme, JT, introduced in Section 2.2.3 requires 0.5μs timing accuracy in collaboration between base stations, which is the tightest constraint. However, it is easier to cope with synchronization challenges in C-RAN compared to traditional RAN due to the fact that BBUs are co-located in the BBU Pool.

2. According to [20], regardless of the delay caused by the cable length, round trip delay of user data may not exceed 5 μs, measured with the accuracy of $\pm 16.276ns$ on each link or hop [33].

3. The sub-frame processing delay on a link between RRHs and BBU should be kept below 1 ms, in order to meet HARQ requirements. Due to the delay requirements of HARQ mechanism, generally maximum distance between RRH and BBU must not exceed 20-40 km [20].

Recommendations on transport network capacity can be found in Section 2.4.

2.3.2 BBU cooperation, interconnection and clustering

Cooperation between base stations is needed to support CoMP in terms of sharing the user data, scheduling at the base station and handling channel feedback information to deal with interference.

Co-location of many BBUs requires special security and resilience mechanisms. Solutions enabling connection of BBUs shall be reliable, support high bandwidth and low latency, low cost with a flexible topology interconnecting RRHs. Thus, C-RAN must provide a reliability that is better or comparable to traditional optical networks like Synchronous Digital Hierarchy (SDH), which achieved high reliability due to their ring topology. Mechanisms like fiber ring network protection can be used.

Cells should be optimally clustered to be assigned to one BBU Pool, in order to achieve statistical multiplexing gain, facilitate CoMP, but also to prevent the BBU Pool and the transport network from overloading. One BBU Pool should support cells from different areas such as office, residential or commercial. After analyzing interferences a beneficial assignment of cells to one BBU Pool can be chosen.

To achieve optimal energy savings of the C-RAN, base stations need to be chosen in a way that will optimize the number of active RRHs/BBU units within the BBU Pool. Proper RRH aggregation and assignment to one BBU Pool can also facilitate CoMP [53].

To achieve optimal throughput on the cell edges cooperative transmission/reception schemes are needed to deal with large Inter Cell Interference (ICI), improving spectrum efficiency. The resource sharing algorithms have been developed by the research community. They need to be combined with an algorithm clustering the cells to reduce scheduling complexity. Therefore, the well-designed scheduler in C-RAN also has an impact on the spectrum efficiency [26].

In [40] Namba *et al.* propose an architecture of Colony RAN that can dynamically change the connections of BBUs and RRHs in respect to traffic demand. Semi-static and adaptive BBU-RRH switching schemes for C-RAN are presented and evaluated in [66], where it was proved that the number of BBUs can be reduced by 26% and 47% for semi-static and adaptive schemes, respectively, compared with the static assignment.

2.3.3 Virtualization technique

A virtualization technique needs to be proposed to distribute or group processing between virtual base station entities and sharing of resources among multiple operators. Any processing algorithm should be expected to work real time - dynamic processing capacity allocation is necessary to deal with a dynamically changing cell load.

Virtualization and cloud computing techniques for IT applications are well defined and developed. However, C-RAN application poses different requirements on cloud infrastructure than cloud computing. Table 2.3 compares cloud computing and C-RAN requirements on cloud infrastructure. More on virtualization solutions can be found in Section 2.7.

Table 2.3: Requirements for cloud computing and C-RAN applications [54]

	IT - Cloud computing	Telecom - Cloud RAN
Client/base station data rate	Mbps range, bursty, low activity	Gbps range, constant stream
Latency and jitter	Tens of ms	< 0.5 ms, jitter in ns range
Life time of information	Long (content data)	Extremely short (data symbols and received samples)
Allowed recovery time	s range (sometimes hours)	ms range to avoid network outage
Number of clients per centralized location	Thousands, even millions	Tens, maybe hundreds

2.4 Transport network techniques

This section presents the technical solutions enabling C-RAN by discussing on transport network, covering physical layer architecture, physical medium, possible transport network standards and devices needed to support or facilitate deployments. Moreover, IQ compression techniques are listed and compared.

As introduced in Section 2.3, a C-RAN solution imposes a considerable overhead on the transport network. In this section, a number of transport network capacity issues are addressed, evaluating the internal architecture of C-RAN and the physical medium in Section 2.4.1 as well as transport layer solutions that could support C-RAN in Section 2.4.2. An important consideration is to apply IQ compression/decompression between RRH and BBU. Currently available solutions are listed in Section 2.4.4.

The main focus of this section is on fronthaul transport network, as this is characteristic for C-RAN. Considerations on backhaul network can be found in, e.g., [67]. The choice of the solution for the particular mobile network operator depends on whether C-RAN is deployed from scratch as green field deployment or introduced on top of existing infrastructure. More on deployment scenarios can be found in Section 2.8.

2.4.1 Physical layer architecture and physical medium

2.4.1.1 PHY layer architecture in C-RAN

There are two approaches on how to split base station functions between RRH and BBU within C-RAN in order to reduce transport network overhead.

In the fully centralized solution, L1, L2 and L3 functionalities reside in the BBU Pool, as shown in Figure 2.8a. This solution intrinsically generates high bandwidth IQ data transmission between RRH and BBU.

In partially centralized solution, shown in Figure 2.8b, L1 processing is co-located with the RRH, thus reducing the burden in terms of bandwidth on the optical transport links, as the demodulated signal occupies 20 - 50 times less bandwidth [20] than the modulated one. This solution is however less optimal because resource sharing is considerably reduced and advanced features such as CoMP cannot be efficiently supported. CoMP benefits from processing the signal on L1, L2 and L3 in one BBU Pool instead of in several base stations [20]. Therefore a fully centralized solution is more optimal. Other solutions, in between the two discussed above, have also been proposed, where only some specific functions of L1 processing are co-located with the RRH, e.g., L1 pre-processing of cell/sector specific functions, and most of L1 is left in the BBU [68].

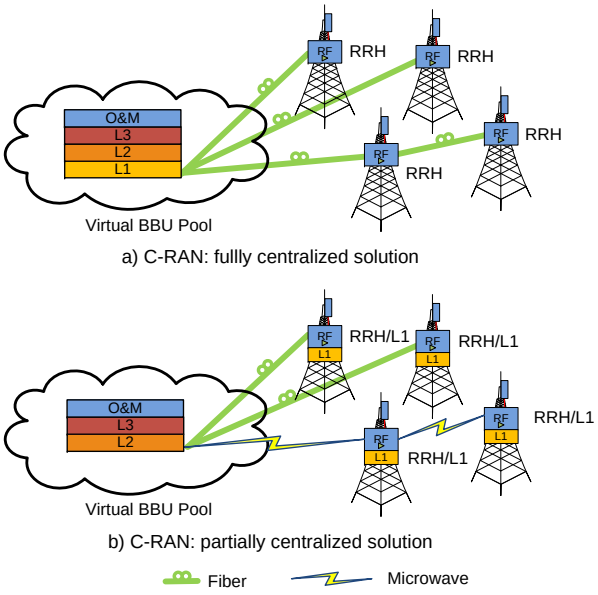


Figure 2.8: C-RAN architecture can be either fully or partially centralized depending on L1 baseband processing module location.

2.4.1.2 Physical medium

As presented in [22], only 35% of base stations were forecasted to be connected through fiber, and 55% by wireless technologies, the remaining 10% by copper on a global scale in 2014. However, the global share of fiber connections is growing. In North America the highest percentage of backhaul connections were forecasted to be done over fiber - 62.5% in 2014 [69].

Fiber links allow huge transport capacity, supporting up to tens of Gbps per channel. 40 Gbps per channel is now commercially available, while future systems will be using 100 Gbps modules and higher, when their price and maturity will become more attractive [20].

Typical microwave solutions offer from 10 Mbps-100 Mbps up to 1 Gbps range [70], the latter available only for a short range (up to 1.5 km) [69]. Therefore 3:1 compression would allow 2.5 Gbps data to be sent over such 1 Gbps link. In [71] Ghebretensae *et al.* propose to use E-band microwave transmission in (70/80 GHz) between BBU Pool and RRH. They proved that E-band microwave transmission can provide Gbps capacity, using equipment currently available commercially (2012) on the distance limited to 1-2 km to assure 99.999% link availability and 5-7 km when this requirement is relaxed to 99.9% availability. In the laboratory setup they have achieved 2.5 Gbps on microwave CPRI links. This supports delivering 60 Mbps to the end user LTE equipment. E-BLINK [72] is an exemplary company providing wireless fronthaul.

For small cells deployment, Wi-Fi is seen as a possible solution for wireless backhauling [67]. Therefore, using the same solutions, Wi-Fi can potentially be used for fronthauling. The latest Wi-Fi standard, IEEE 802.11ad, can achieve the maximum theoretical throughput of 7 Gbps. However, the solution is not available on the market yet (2013).

The solution based on copper links is not taken into account for C-RAN, as Digital Subscriber Line (DSL) based access can offer only up to 10-100 Mbps.

To conclude, full C-RAN deployment is currently only possible with fiber links between RRH and BBU Pool. In case C-RAN is deployed in a partially centralized architecture, or when compression is applied, microwave can be considered as a transport medium between RRHs and BBU Pool.

2.4.2 Transport network

As fiber is the most prominent solution for the physical medium, its availability for the network operator needs to be taken into account choosing the optimal transport network solution. Moreover, operators may want to reuse their existing deployments. Various transport network solutions are presented in Figure 2.9 and discussed below [20].

2.4.2.1 Point to point fiber

Point to point fiber is a preferred solution for a BBU Pool with less than 10 macro base stations [20], due to capacity requirements. Dark fiber can be used with low cost,

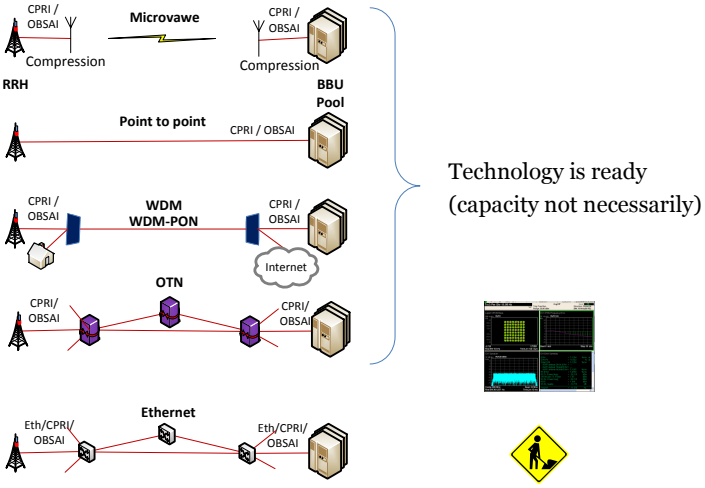


Figure 2.9: Possible fronthaul transport solutions

because no additional optical transport network equipment is needed. On the other hand, this solution consumes significant fiber resources, therefore network extensibility is a challenge. New protection mechanisms are required in case of failure, as well as additional mechanisms to implement O&M are needed. However, those challenges can be answered. CPRI products are offering 1+1 backup/ring topology protection features. If fiber is deployed with physical ring topology it offers resiliency similar to SDH. O&M capabilities can be introduced in CPRI.

2.4.2.2 WDM/OTN

Wavelength-division multiplexing (WDM)/Optical Transport Network (OTN) solutions are suitable for macro cellular base station systems with limited fiber resources, especially in the access ring. The solution improves the bandwidth on BBU-RRH link, as 40-80 optical wavelength can be transmitted in a single optical fiber, therefore with 10 Gbps large number of cascading RRH can be supported, reducing the demand on dark fiber. On the other hand, high cost of upgrade to WDM/OTN need to be covered. However, as the span on fronthaul network does not exceed tens of kilometers, equipment can be cheaper than in long distance backbone networks. Usage of plain WDM CPRI transceivers was discussed and their performance was evaluated in [73]. [23] applies WDM in their vision of C-RAN transport network.

In [74] Ponzini *et al.* describe the concept of non-hierarchical WDM-based access for

C-RAN. The authors have proven that WDM technologies can more efficiently support clustered base station deployments offering improved flexibility in term of network transparency and costs. Using that concept already deployed fibers, such as Passive Optical Networks (PONs) or metro rings, can be reused to carry any type of traffic, including CPRI, on a common fiber infrastructure. By establishing virtual P2P WDM links up to 48 bidirectional CPRI links per fiber can be supported.

For scarce fiber availability ZTE proposes enhanced fiber connection or xWDM/OTN [65]. Coarse WDM is suitable to be used for TD-SCDMA, while Dense WDM for LTE, due to capacity requirements.

OTN is a standard proposed to provide a way of supervising client's signals, assure reliability compared with Synchronous Optical NETworking (SONET)/SDH network as well as achieve carrier grade of service. It efficiently supports SONET/SDH as well as Ethernet and CPRI. CPRI can be transported over OTN over low level Optical channel Data Unit (ODU)k containers as described in ITU-T G.709/Y.1331 [75], [76].

2.4.2.3 Unified Fixed and Mobile access

Unified Fixed and Mobile access, like UniPON, based on Coarse WDM, combines fixed broadband and mobile access network. UniPON provides both PON services and CPRI transmission. It is suitable for indoor coverage deployment, offers 14 different wavelengths per optical cable, reducing overall cost as a result of sharing. However, it should be designed to be competitive in cost. Such a WDM-OFDMA UniPON architecture is proposed and examined in [77], and a second one, based on WDM-PON in [71]. In [71], referenced also in Section 2.4.1.2, Ghebretensae *et al.* propose an end-to-end transport network solution based on Dense WDM(-PON) colorless optics, which supports load balancing, auto configuration and path redundancy, while minimizing the network complexity. In [78] Fabrega *et al.* show how to reuse the deployed PON infrastructure for RAN with RRHs. Connections between RRHs and BBUs are separated using very dense WDM, coherent optical OFDM helps to cope with narrow channel spacings.

2.4.2.4 Carrier Ethernet

Carrier Ethernet transport can also be directly applied from RRH towards BBU Pool. In that case, CPRI-to-Ethernet (CPRI2Eth) gateway is needed between RRH and BBU Pool. CPRI2Eth gateway needs to be transparent in terms of delay. It should offer multiplexing capabilities to forward different CPRI streams to be carried by Ethernet to different destinations.

The term Carrier Ethernet refers to two things. The first is the set of services that enable to transport Ethernet frames over different transport technologies. The other one is a solution how to deliver these services, named Carrier Ethernet Transport (CET). Carrier Ethernet, e.g., Provider Backbone Bridge - Traffic Engineering (PBB-TE) is supposed to provide carrier - grade transport solution and leverage the economies of scale of traditional Ethernet [79]. It is defined in IEEE 802.1Qay-2009 standard. It evolved from

IEEE 802.1Q Virtual LAN (VLAN) standard through IEEE 802.1ad Provider Bridges (PB) and IEEE 802.1ah Provider Backbone Bridges (PBB). To achieve Quality of Service (QoS) of Ethernet transport service, traffic engineering is enabled in Carrier Ethernet. PBB-TE uses the set of VLAN IDs to identify specific paths to given MAC address. Therefore a connection-oriented forwarding mode can be introduced. Forwarding information is provided by management plane and therefore predictable behavior on predefined paths can be assured. Carrier Ethernet ensures 99.999% service availability. Up to 16 million customers can be supported which removes scalability problem of PBB-TE predecessor [80]. Carrier Ethernet grade of service can also be assured by using MPLS Transport Profile (MPLS-TP). Technologies are very similar, although PBB-TE is based on Ehternet and MPLS-TP on Multiprotocol Label Switching (MPLS).

The main challenge in using packet passed Ethernet in the fronthaul is to meet the strict requirements on synchronization, syntonization and delay. Synchronization refers to phase and syntonization to the frequency alignment, respectively. Base stations need to be phase and frequency aligned in order to, e.g., switch between uplink and downlink in the right moment and to stay within their allocated spectrum. For LTE-A frequency accuracy needs to stay within $\pm 50ppb$ (for a wide area base station) [6.5 in [81]] while phase accuracy of $\pm 1.5\mu s$ is required for cell with radius $\leq 3km$ [82]. Time-Sensitive Networking (TSN) features help to achieve delay and synchronization requirements. More information about them can be found in Section 4.7, while the whole Chapter 4 gives a deep overview of challenges and solutions for Carrier Ethernet-based fronthaul.

Altiostar is an exemplary company providing Ethernet fronthaul [83].

2.4.3 Transport network equipment

This Section presents examples of network equipment that has been developed for usage in C-RAN architecture.

2.4.3.1 CPRI2Ethernet gateway

If Ethernet is chosen as a transport network standard, CPRI2Eth gateway is needed to map CPRI data to Ethernet packets, close to or at the interface of RRH towards BBU Pool. Patents on such a solutions have been filed, see for example, [84].

2.4.3.2 IQ data routing switch

China Mobile Research Institute developed a large scale BBU Pool supporting more than 1000 carriers in 2011. The key enabler of this demonstration was a IQ data routing switch [20]. It is based on a Fat-Tree architecture of Dynamic Circuit Network (DCN) technology. In Fat-Tree topology multiple root nodes are connected to separate trees. That ensures high reliability and an easy solution to implement load balancing between BBUs. China Mobile has achieved real time processing and link load balancing. In addition, resource management platform has been implemented.

2.4.3.3 CPRI MUX

CPRI mux is a device that aggregates traffic from various radios and encapsulates it for transport over a minimum number of optical interfaces. It can also implement IQ compression/decompression and have optical interfaces: for Coarse WDM and/or Dense WDM. BBU Pool will be demultiplexing the signals multiplexed by the CPRI mux [22].

2.4.3.4 x2OTN GATEWAY

If OTN is chosen as a transport network solution, then CPRI/OBSAI to OTN gateway is needed to map signals from two standards. Altera has a Soft Silicon OTN processor that can map any client into ODU container [85]. The work was started by TPACK. Performance of CPRI and OBSAI over OTN transport network has been proven in [86] for e.g., C-RAN application.

2.4.4 IQ Compression schemes and solutions

In C-RAN the expected data rate at the fronthaul link can be 12 to 55 times higher compared to data rate on the radio interface, depending on CPRI IQ sample width and modulation. RRHs transmit raw IQ samples towards BBU cloud, therefore, an efficient compression schemes are needed to optimize such a huge bandwidth transmission over capacity-constrained links. Potential solutions could be to reduce signal sampling rate, use non-linear quantization, frequency sub-carrier compression or IQ data compression [20]. Techniques can be mixed and a chosen scheme is a trade-off between achievable compression ratio, algorithm and design complexity, computational delay and the signal distortion it introduces as well as power consumption, as shown in Figure 2.10. The following techniques can be used to achieve IQ compression.

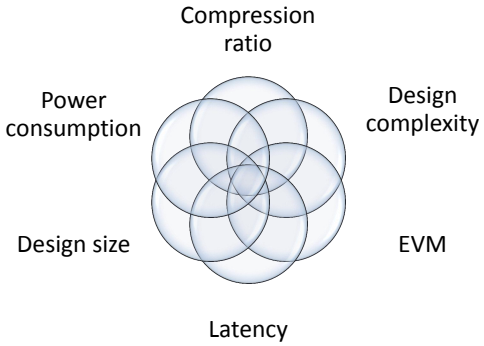


Figure 2.10: Factors between which a trade off needs to be reached choosing an IQ compression scheme.

Reducing signal sampling rate is a low complex solution having minimal impact on protocols, improves compression up to 66% with some performance degradation [20].

By applying **non-linear quantization**, more quantization levels are specified for the region in magnitude where more values are likely to be present. This solution improves Quantization SNR (QSNR). Mature, logarithmic encoding algorithms, like μ-Law or A-law are available to specify the step size. Compression efficiency up to 53% can be achieved. This method creates additional Ir interface complexity (interface between RRH and BBU) [20].

IQ data compression can be done using e.g., Digital Automatic Gain Control (DAGC) [20], [87]. This technique is based on reducing the signal's dynamic range by normalizing the power of each symbol to the average power reference, therefore reducing the signal dynamic range. This method affects Signal-to-noise ratio (SNR) and Error Vector Magnitude (EVM) deteriorates in DL. Potential high compression rate can be achieved, however the method has a high complexity and no mature algorithms are available.

One example of a frequency domain scheme is to perform **sub carrier compression**. Implementing the FFT/Inverse FFT (IFFT) blocks in the RRH allows 40% reduction of Ir interface load. It can be easily performed in DL, however RACH processing is a big challenge. This frequency domain compression increases IQ mapping and system complexity. It also requires costly devices, more storage and larger FPGA processing capacity [20]. On top of that, it limits the benefits of sharing the equipment in C-RAN, as L1 processing needs to be assigned to one RRH. Several patents have been filed for this type of compression schemes.

In [88] Grieger *et al.* present design criteria for frequency domain compression algorithms for LTE-A systems which are then evaluated in large scale urban filed trials. Performance of JD under limited backhaul rates was observed. The authors proved that a Gaussian compression codebook achieves good performance for the compression of OFDM signals. The performance can be improved using Frequency Domain AGC (FDAGC) or decorrelation of antenna signals. However, field tests showed very limited gains for the observed setups.

Samardzija *et al.* from Bell Laboratories propose an algorithm [89] which reduces transmission data rates. It removes redundancies in the spectral domain, performs block scaling, and uses a non-uniform quantizer. It keeps EVM below 8% (3GPP requirement for 64 QAM, as stated in [81]) for 17% of relative transmission data rate (compression ratio defined as transmission rate achieved after compression to the original one). The algorithm presented by Guo *et al.* [90], which authors are also associated with Alcatel-Lucent Bell Labs removes redundancies in spectral domain, preforms block scaling, and uses non-uniform quantizer. EVM stays within 3GPP requirements in simulations for 30% compression ratio. TD-LTE demo test results showed no performance loss for 50% compression ratio.

Alcatel-Lucent Bell Labs' compression algorithm reduces LTE traffic carried over CPRI interface from 18 Gbps to 8 Gbps [22], achieving a 44% compression ratio.

The solution discussed in [91] adapts to the dynamic range of the signal, removes frequency redundancy and performs IQ compression creating 10.5 effective bits out of 12

bits of data. This method allows 50% to 25% of compression ratio introducing 0.5% [1] to 8% of EVM and latency below $1\mu s$ for LTE signal.

Lorca *et al.* from Telefonica I + D in [92] propose a lossless compression technique where actual compression ratios depend upon the network load. For downlink direction, the algorithm removes redundancies in the frequency domain. Secondly, the amount of control data is reduced to minimum sending only the necessary information to reconstruct control signals at RRH. Moreover, a special constellation coding is used to reduce number of bits needed to represent constellation symbols for QPSK, 16QAM and 64QAM modulations. For uplink direction user detection is used to transmit only occupied carriers. Compression ratio of 33% is achieved at full cell load. Compression ratio up to 6.6% are achieved for 20% cell load.

Park *et al.* [93] propose a robust, distributed compression scheme applicable for UL transmission, which they combine with an efficient base station selection algorithm. Their current work focuses on implementing layered compression strategy as well as joint decompression and decoding. Results in terms of compression ratio and EVM are not available.

Table 2.4 summarizes and compares various compression methods discussed in this section. Compression of 33% is achieved by all the algorithms for which the ratio was available. The best result, where the algorithm is known, is achieved by [89] and by [92] under small network load.

To conclude, in order not to lose the cost benefit of BBU Pooling for renting a transport network, mobile network operator needs to either own substantial amount of fiber or use an IQ compression scheme. Moreover, the cost of the optical high speed module must stay comparable to traditional SDH transport equipment in order to make C-RAN economically attractive.

2.5 RRH development

This section presents requirements and solutions for RRH that are compatible with C-RAN. The existing RRHs are expected to work in a fully centralized C-RAN architecture in a plug-and-play manner. In case of partially centralized C-RAN architecture L1 needs to be incorporated in RRH.

The biggest difference between RRHs deployed for C-RAN compared to previous solutions is that in C-RAN transmission the signal occurs over many kilometers, while in the latter architecture this distance is shorter, typically up to few kilometers. Therefore the additional delay caused by increased transmission distance needs to be monitored.

In addition, the higher bit rates need to be supported. In order to transport 10 Gbps CPRI rate, the maximum CPRI line bit rate option 8, i.e., 10.1376 Gbps needs to be deployed, which is supported so far by standard CPRI v 6.0 [33]. Additional upgrade of the standard is needed to accommodate more traffic, at least 16 Gbps to fully serve a 3

[1] equivalent to test equipment

Table 2.4: Comparison of IQ compression methods. Compression ratio 33% corresponds to 3:1

Method	Techniques applied	Compression ratio	EVM
[22]	Not available	44%	Not available
[89]	removing redundancies in spectral domain	28%	3%
	preforming block scaling	23%	4%
	usage of non-uniform quantizer	17%	8%
[90]	removing redundancies in spectral domain	52%	$<$ 1.4%
	preforming block scaling	39%	$<$ 1.5%
	usage of non-uniform quantizer	30%	$<$ 2.5%
[91]	adaption of dynamic range of the signal	50%	0.5%
	removal of frequency redundancy	33%	3%
	IQ compression	25%	8%
[92]	removal of frequency redundancy optimized control information transmission IQ compression user detection	33% (100% cell load) 7% (20% cell load)	Not available
[93]	self-defined robust method performed jointly with base station selection algorithm	Not available	Not available

sector 20 MHz LTE macro cell with 4x2 MIMO [22], see Table 2.2. Existing standards - CPRI and OBSAI can support connections between the BBU Pool and RRHs in C-RAN. Moreover, NGMN in [94] envisions ORI as a future candidate protocol. However, as the nature of the interface between RRH and BBU is changing with an introduction of C-RAN, the existing protocols may need to be redefined in order to be optimized for high volume transmission over long distances.

Alcatel-Lucent is offering a lightRadio solution for C-RAN [22]. It uses a multiband, multistandard active antenna array, with MIMO and passive antenna array support. Alcatel-Lucent is working towards two multiband radio heads (one for high and one for low bands). Built-in digital modules are used for baseband processing. For C-RAN L1, L2 and L3 are separated from radio functions.

In 2012, Ericsson announced the first CPRI over microwave connection implementation [95], which is interesting for operators considering the deployment of a partially centralized C-RAN architecture.

2.6 Synchronized BBU Implementation

This section provides considerations on possible BBU implementation. The advantages and disadvantages of different processors types that can be used in C-RAN are discussed.

The interconnection between BBUs is required to work with low latency, high speed, high reliability and real time transmission of 10 Gbps. Furthermore, it needs to support CoMP, dynamic carrier scheduling, 1+1 failure protection and offer high scalability. Dynamic carrier scheduling implemented within the BBU Pool enhances redundancy of BBU and increases reliability.

The BBU Pool needs to support 100 base stations for a medium sized urban network (coverage 5x5 km), 1000 base stations for 15x15 km [20]. In addition, it is beneficial when BBU has the intelligence to support additional services like Content Distribution Network (CDN), Distributed Service Network (DSN) and Deep Packet Inspection (DPI) [25].

Virtualization of base station resources is needed to hide the physical characteristics of the BBU Pool and enable dynamic resource allocation.

There are also challenges for real time virtualized base station in centralized BBU Pool, like high performance low-power signal processing, real time signal processing, BBU interconnection as well as between chips in a BBU, BBUs in a physical rack and between racks.

Optimal pooling of BBU resources in needed in C-RAN. In [42] Bhaumik *et al.* propose resource pooling scheme to minimize the number of required compute resources. The resource pooling time scale is of the order of several minutes, however, it can be expected it can be done with finer granularity further optimizing the results.

2.6.1 Current multi-standard open platform base station solutions

Operators need to support multiple standards, therefore multi-mode base stations are a natural choice. They can be deployed using either pluggable or software reconfigurable processing boards for different standards [20].

By separating the hardware and software, using e.g., SDR technology, different wireless standards and various services can be introduced smoothly. Currently base stations are built on proprietary platforms (vertical solution). C-RAN is intended to be build on open platforms in order to relief mobile operators from managing multiple, often non-compatible platforms. C-RAN provides also higher flexibility in network upgrades and fosters the creation of innovative applications and services.

2.6.2 Processors

Nowadays, Field-Programmable Gate Arrays (FPGAs) and embedded Digital Signal Processor (DSP) are used for wireless systems. However, the improvement in the processing power of General Purpose Processor (GPP) used in IT is giving the possibility to bring IT and telecom worlds together and use flexible GPP-based signal processors.

DSPs are developed to be specially optimized for real-time signal processing. They are powerful and use multicore (3-6) technology with improved processing capacity. What is important for C-RAN, a real time OS running on DSP facilitates virtualization of processing resources in a real time manner. However, there is no guarantee of backwards compatibility between solutions from different, or even from the same manufacturer, as they are built on generally proprietary platforms.

Texas Instruments [26] favors the usage of specialized wireless System on a Chip (SoC), providing arguments that SoC consumes one-tenth of the power consumed by a typical server chip, and has wireless accelerators and signal processing specialization. Considerations about power consumption of signal processors are essential to achieve reduction in power consumption for C-RAN architecture compared to the traditional RAN. In addition, for the same processing power, a DSP solution will also have a lower price compared to GPP. In [96] Wei *et al.* present an implementation of SDR system on an ARM Cortex-A9 processor that meets the real-time requirements of communication system. As SDR technology further enables to benefit from C-RAN this is an important proof of concept.

GPPs are getting more and more popular for wireless signal processing applications. The usage of GPP is facilitated by muli-core processing, single-instruction multiple data, low latency off-chip system memory and large on-chip caches. They also ensure backward compatibility, which makes it possible to smoothly upgrade the BBU. Multiple OS's with real-time capability allow virtualization of base station signal processing.

China Mobile Research Institute proved that commercial IT servers are capable of performing signal processing in a timely manner. Intel is providing the processors for both C-RAN and traditional RAN [25]. More on Intel GPP solutions for DSP can be found in [97]. In [98], Kai *et al.* present a prototype of a TD-LTE eNB using a GPP.

It did not meet real-time requirements of LTE system, which is of great concern when using general processors for telecommunication applications. It used 6.587 ms for UL processing, with turbo decoding and FFT taking most of it and 1.225 ms for downlink (DL) processing, with IFFT and turbo coding being again the most time consuming. However, this system was based on a single core, and multi-core implementation with 4 cores should make the latency fall within the required limits. Another approach to reach the requirements is to optimize the turbo decoder as described in [99], where Zhang *et al.* prove that using multiple threads and a smart implementation, 3GPP requirements can be met. De-Rate Matching and demodulation have been optimized for GPP used for LTE in [100]. In [101] Kaitz *et al.* propose to introduce a dedicated co-processor optimized for wireless and responsible for critical and computation intensive tasks. This optimizes power consumption at the cost of decreased flexibility. They have considered different CPU partitioning approaches for LTE-A case.

The issue of real-time timing control and synchronization for SDR has been addressed in [102]. A real-time and high precision clock source is designed on a GPP-based SDR platforms and users are synchronized utilizing Round-Trip Delay (RTD) algorithm. The mechanism is experimentally validated.

Table 2.5 summarizes the characteristics of DSP and GPP.

Table 2.5: DSP and GPP processors

	DSP	GPP
Flexibility	dedicated solution	general purpose
Vendor compatibility	vendor specific, proprietary	higher compatibility between vendors
Backward compatibility	limited	assured
Power consumption	lower	higher
Real-time processing	optimized, achieved	only possible with high power hardware
Virtualization of BBU	possible	possible

2.7 Virtualization

In order to implement Centralized RAN, BBU hardware needs to be co-located. However, in order to deploy RAN in the cloud - Cloud RAN - virtualization is needed. Although virtualization is not the focus point of this thesis, it is briefly presented in this section for the completeness of the C-RAN introduction, as it is an important foundation of C-RAN.

Virtualization enables decoupling network functions, like baseband processing, from network hardware in order to reduce the network cost and to enable flexible deployments of services [103]. Legacy hardware can be used for new applications. Moreover, utilization of the hardware can be increased. NFV and SDN are important concepts that help to implement virtualization of baseband resources in C-RAN. The common goals of NFV and SDN are to allow network services to be automatically deployed and programmed.

NFV is an architecture framework currently under standardization by ETSI. The framework relies on three different levels: Network Functions Virtualization Infrastructure, Virtual Network Functions (applications) and NFV Orchestrator (control and management functions).

SDN [104] is a concept related to NFV. SDN decouples data from the control plane. In that way the control plane can be directly programmable while the underlying physical infrastructure can be abstracted from applications and services. In can bring greater flexibility, fine-grained control, and ease-of-use to networking. OpenFlow [105] is a protocol commonly used to organize communications between between network devices (e.g. switches) and the controller, on the so-called southbound interface. Northbound interface refers to the communication between the controller and the network applications. OpenDaylight [106] is an open source project whose goal is to accelerate the adoption of SDN.

SDN is not required to implement NFV and vice-versa. NFV alone allows deploying infrastructure services on open hardware. However, NFV is often combined with SDN, as SDN is beneficial to the implementation and management of an NFV infrastructure [103]. OpenNFV [107] is an exemplary NFV platform managing various components. This platform integrates orchestration and management functions that run on top of Virtual Network Functions. Openstack [108] controls Compute Virtualization Control, Storage Virtualization Control and Network Virtualization Control. OpenDaylight (an SDN tool) is used for Network Virtualization Control, however, other mechanisms could be used instead of SDN for the networking part.

Altiostar is a sample company providing virtualized RAN based on an open source NFV platform [83]. Moreover, Wind River [109] offers an NFV software platform.

Nikaein *et al.* in [110] presents a virtualized RAN implemented with OpenStack and Heat stack. Heat implements an orchestration engine to manage multiple composite cloud applications: OpenAirInterface [111] eNB, Home Subscriber Server (HSS) and Evolved Packet Core (EPC).

2.8 Likely deployment Scenarios

C-RAN is intended to be an alternative delivery of cellular standards, like UMTS, LTE, LTE-A and beyond. It is a RAN deployment applicable to most typical scenarios, like macro-, micro- and picocell, as well as for indoor coverage. This section elaborates on likely deployment scenarios for C-RAN including green field deployments, i.e., estab-

lishing the network from scratch, as well as deployment of additional cells for boosting the capacity of an existing network. Moreover, different stages of C-RAN deployment to leverage its full potential are listed.

It is advised to deploy C-RAN in metropolitan area to benefit from statistical multiplexing gain, as users are moving through the day, but still remain within the maximum distance (resulting from propagation and processing delay, up to 40 km) between RRH and BBU. However, a metropolitan area might be served by a few BBU Pools.

2.8.1 Green field deployment

In case of green field deployment, RRH and BBU Pool placement need to be arranged according to network planning. Physical medium and transport solution can be designed according to C-RAN specific requirements.

In author's previous work [12] the most beneficial C-RAN deployments are evaluated. For the analyzed traffic model, to maximize statistical multiplexing gain it is advisable to serve 20-30% of office base stations and 70-80% of residential base stations in one BBU Pool. Both analytical and simulation - based approach confirm the results.

The analysis on the cost of deployments from the same work shows that in order to minimize TCO a ratio of cost of one BBU to the cost of one kilometer of fiber deployment should be above 3. The ratio is smaller looking at smaller (100 km^2) areas compared to larger (400 km^2) areas. Therefore, C-RAN is more promising for small scale deployments for urban areas with densely placed cells.

2.8.2 C-RAN for capacity boosting

Small cells are a likely scenario for RRHs and C-RAN. Release 12 of mobile standards addresses enhancement of small cell deployment [112], as adding new cells is the most promising way to increase network capacity. In [70] authors envision that small cells enhancements will be deployed with and without macro coverage, sparsely or densely, outdoor and indoor, being connected through ideal and non-ideal backhaul. Frequencies will be separately assigned to macro- and small cells. C-RAN fits into these target scenarios. It also fulfills the requirements for small cells enhancements, supporting both operator and user deployed cells, SON mechanisms as well as co-existence and networking between different RATs.

In mobile networks within an underlying macro cell many small cells can be deployed to boosts network capacity and quality in homes, offices and public spaces. When a user will move out of small cell coverage, he will change the cell to the macro cell. In order to support such an architecture, a coordination is required between macro- and small cells. The deployment of small cells with C-RAN architecture may reduce signaling resources as they are supported by one BBU pool, not many base stations. To deploy C-RAN for capacity improvement, some of the existing BBUs can be moved to the BBU Pool. RRHs can remain in the same location, and additional ones can be added. Various possibilities

of capacity improvement deployment scenarios are listed below [68]. The combination of mentioned solutions is also possible.

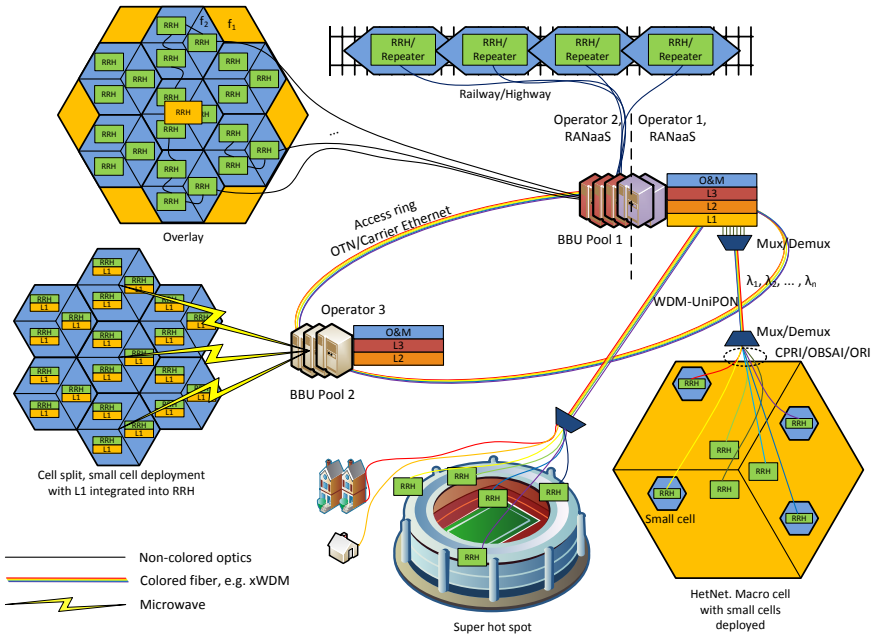


Figure 2.11: C-RAN deployment scenarios.

a) HetNets. Existing BBUs of macro Base Stations can be replaced by BBU Pool and additional RRHs can be deployed to form small cells.

b) Cell split. Existing macro cells can be split into smaller ones increasing the system capacity. Interference management techniques are needed as all the cells will operate at the same frequency. As explained in Section 2.2.3, C-RAN can enhance cooperative techniques like CoMP and eICIC. This scenario can also be used to provide indoor coverage by deploying RRHs on each floor of the building or group of offices offering high capacity. However, in this scenario Wi-Fi can be a cheaper solution, if users will have Wi-Fi connection in their devices switched on, enabling offload from cellular network to Wi-Fi.

c) Overlay. Additional frequency band or a new cellular standard can be introduced to boost system capacity. In Figure 2.11 one RRH provides coverage in frequency f_1.

Additional RRHs operating on frequency f_2 provide overlay coverage. Efficient interference management techniques like CoMP and eICIC are needed in this scenario, as many RRH operate at the same frequency f_2.

d) Super hot spots, e.g., stadium, transportation hub. It is a scenario where many users are present in one location. Small cells are needed to assure the capacity, as well as provide the coverage in complex scenery, e.g., with balconies, ramps, etc. The density of users is high, therefore it is crucial to efficiently support interference management schemes like CoMP and eICIC.

e) Railway/highway. Users are moving with a fast speed in this scenario, therefore BBU Pool shall handle frequent handovers faster than traditional RAN.

Figure 2.11 summarizes C-RAN transport solutions and physical layer architecture discussed in this chapter. Moreover, a possibility of sharing BBU Pool and rent RANaaS is emphasized. For a particular network operator the choice of physical medium and transport network depends on whether an existing infrastructure is already deployed.

2.8.3 Different stages of deployment

The path towards complete deployment of C-RAN can be paved through following stages [113].

1. Centralized RAN, where baseband units are deployed centrally supporting many RRHs. However, resources are not pooled, nor virtualized.

2. Cloud RAN

   - Phase 1, where baseband resources are pooled. Baseband processing is done using specialized baseband chip - DSPs,

   - Phase 2, where resources are virtualized, using GPP, thereby leveraging full benefits of C-RAN. Sometimes this deployment is referred to as V-RAN standing for Virtualized-RAN.

2.9 Ongoing work

This section introduces projects focused on C-RAN definition and development. Moreover, the survey on field trials and developed prototypes as well as the announcement of first commercial deployment is presented.

2.9.1 Joint effort

Both academic and industrial communities are focusing their attention on C-RAN in a number of projects. China Mobile has invited industrial partners to sign Memorandum of Understanding (MoU) on C-RAN. The companies mentioned below have already signed a MoU with China Mobile Research Institute, and therefore engaged to work on novel C-RAN architectures: ZTE, IBM, Huawei, Intel, Orange, Chuanhua Telecom, Alcatel-Lucent, Datang Mobile, Ericsson, Nokia Siemens Networks and recently (February 2013) ASOCS.

The Next Generation Mobile Networks (NGMN) alliance has proposed requirements and solutions for a new RAN implementation in the project "Project Centralized processing, Collaborative Radio, Real-Time Computing, Clean RAN System (P-CRAN)" [114]. One of the project outcomes is a description of use cases for C-RAN and suggestions for solutions on building and implementing C-RAN [68].

OpenAirInterface [111] is a software alliance working towards an open source software running on general purpose processors. Currently (February 2016) is offers a subset of Release 10 LTE for UE, eNB, Mobility Management Entity (MME), HSS, Serving SAE Gateway (SAE-GW) and Packet Data Network Gateway (PDN-GW) on standard Linux-based computing equipment.

Three projects sponsored by the Seventh Framework Programme (FP7) for Research of the European Commission have been running since November 2012. The "Mobile Cloud Networking" (MCN) project [115] evaluated and seized the opportunities that cloud computing can bring to mobile networks. It was the biggest out of FP7 projects in terms of financial resources. 19 partners worked on decentralized cloud computing infrastructure that provided an end-to-end mobile network architecture from the air interface to the service platforms, using cloud computing paradigm for an on-demand and elastic service. The "High capacity network Architecture with Remote Radio Heads & Parasitic antenna arrays" (HARP) project [116] focused on demonstrating a novel C-RAN architecture based on RRHs and electronically steerable passive antenna radiators (ESPARs), which provide multi-antenna-like functionality with a single RF chain only. The "Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks" (IJOIN) project [117] introduced the novel concept of RAN-as-a-Service (RANaaS) [118], where RAN is flexibly realized on a centralized open IT platform based on a cloud infrastructure. It aimed at integrating small cells, heterogeneous backhaul and centralized processing. The main scope of the CROWD project [119] were very dense heterogeneous wireless access networks and integrated wireless-wired backhaul networks. The focus was put on SDN, which is relevant for C-RAN. iCirrus [120], an EU Horizon 2020 project, aims to enable a 5G network based on Ethernet transport and switching for fronthaul, backhaul and midhaul (see Section 4.10). ERAN is a nationally founded Danish project that brings together SDN and Ethernet technologies, including TSN features, to provide flexibility and cost reductions in the fronthaul [121]. Table 2.6 summarizes research directions relevant for C-RAN and in

which works they have been addressed up to beginning of 2014.

2.9.2 C-RAN prototype

China Mobile, together with its industry partners - IBM, ZTE, Huawei, Intel, Datang China Mobile, France Telecom Beijing Research Center, Beijing University of Post and Telecom and China Science Institute developed GPP based C-RAN prototype supporting GSM, TD-SCDMA and TD-LTE. The prototype is running on Intel processor-based servers [25]. A commercial IT server processes IQ samples in real time. PCI Express, a high-speed serial computer expansion bus is connected to CPRI/Ir interface converter, which carries the signal towards RRHs. L1, L2 and L3 of GSM and TD-SCDMA as well as L1 TD-LTE are supported. Future plans cover implementing L2 and L3 of TD-LTE and LTE-A features like CoMP [20].

Ericsson Beijing proved their concept of connecting LTE RRH and BBU using WDM-PON and the microwave E-band link, as described in [71]. This proves the novel transport network concept, that can be used for C-RAN. However, the test was done for only 2.5 Gbps connections, while 10 Gbps is desired for C-RAN macro base stations. Moreover, at Ericsson Beijing setup, the joint UL COMP was evaluated in [56]. NEC built OFDMA-based (here WiMAX) C-RAN test-bed with a reconfigurable fronthaul [47].

2.9.3 China Mobile field trial

China Mobile is running C-RAN trials in commercial networks in several cities in China since 2010 [20].

In the GSM trial of C-RAN in Changsha 18 RRHs were connected in daisy-chain with one pair of fiber [20], [65]. By using multi-RRH in one cell, improvement in radio performance and user experience was measured. Reduced inter-site handover delay was achieved, as handover was handled within one BBU Pool.

The trial in Zhuhai City, done on a TD-SCDMA network showed advantages in terms of cost, flexibility and energy saving over a traditional RAN. Dynamic carrier allocation adapted to dynamic load on the network. No change of Key Performance Indicators (KPI) for radio performance was observed. CAPEX and OPEX were reduced by 53% and 30%, respectively for new cell sites compared to traditional RAN. Reduced A/C consumption was observed for C-RAN compared to RAN with RRH. A decrease in base station construction and maintenance cost were also observed. Moreover, base station utilization was improved leading to reduced power consumption [20].

In the field trial in Guangzhou the dual-mode BBU-RRH supported 3G/4G standards. On 12 sites 36 LTE 20 MHz carriers were deployed [49].

2.9.4 First commercial deployments

A number of operators: China Mobile, At&T, Orange, SK Telecom, SoftBank Corp [63] and NTT Docomo [144] are (in 2015) at various stages of C-RAN deployment whit Asia

Table 2.6: Research directions for C-RAN

Research direction	Summary	References
Quantifying multiplexing gains, energy and cost savings	1) Dynamic changes of RRH-BBU Pool assignment as well pooling the resources within a BBU Pool helps maximizing multiplexing gains in C-RAN. 2) Work on evaluating energy and cost savings in C-RAN is ongoing, where a multiplexing gain is one of the factors.	1) [26], [40], [43], [11], [12], [42], [41], [47], [53], [66]; 2) [20], [24], [48], [49], [50]
Quantifying an increase of throughput	It has been analyzed to what extend the cooperative techniques such as ICIC, CoMP and Massive MIMO can be enhanced in C-RAN.	[19], [20], [48], [54], [53], [55], [56], [57], [58], [59]
Wireless fronthaul for C-RAN	Although primary physical medium for C-RAN fronthaul is fiber, there are efforts to make transmission possible through microwave or even, on short distances through Wi-Fi.	[67], [71], [95]
Optical fronthaul for C-RAN	R&D efforts focus on evaluation and optimization of optical transmission employing WDM, OTN, PON and Ethernet.	[20], [22], [23], [65], [71], [73], [74], [77], [78], [85], [86]
IQ compression	In order to reduce the need of a high bandwidth on the fronthaul links, various compression schemes were proposed utilizing signal properties as well as varying network load.	[20], [22], [88], [89], [90], [91], [92], [93]
Moving towards software - virtualization solutions	1) Various works on network, resource and hardware virtualization in wireless communication is relevant for BBU Pool virtualization in C-RAN. By means of 2) NFV and 3) SDN benefits can be further leveraged.	1) [20], [21], [25], [122], [123], [124], [125], [126], [127], [128], [129], [130], [131], [132], [133], [134]; 2) [135], [136], [137], [138]; 3) [139], [140], [141], [142], [119], [143]
Deployment scenarios	Literature summarizes considerations on deployment scenarios covering the optimal architectures for the given fiber resources as well as possibilities of deployments to boost the capacity of the network. Moreover, an analysis has been done on how to maximize the multiplexing gains by grouping cells with a given traffic profiles in a BBU Pool.	[22], [30], [12], [68]

leading the effort. More operators, like Vodafone Hutchison Australia are planning the deployments in 2018-2020 time frame.

2.10 Future directions

As it follows from the previous sections, CPRI and OBSAI are simple protocols that served well for short range separation between RRH and BBU. Typically the distance between RRH and BBU would be from the basement to the rooftop and a dedicated point to point fiber connection could be easily established for carrying IQ data. With the introduction of C-RAN, in order to benefit from the multiplexing gain coming from users' daily mobility between office and residential sites it is expected that the BBU pool would often span the metropolitan area serving tens of cells. It will thereby require 100s of Gbps connections, which are often not available. The industry and academia are working on defining a new functional split between RRH and BBU to address this issue. Moreover, an optimized transport is needed to carry the data for the new functional split.

Sections below provide more details about those directions. In the meanwhile the CPRI consortium works on newer editions of the standard, with the latest one (6.1) supporting up to 12 Gbps [145]. Moreover, observations on centralization and decentralization trends in base station architecture are provided.

2.10.1 Next Generation Fronthaul Interface (NGFI), functional split

With the current functional split, all L1-L3 baseband processing is executed by the BBU making an RRH a relatively simple and cheap device. However, the following shortcomings of CPRI motivate work on changing the functional split giving up some of the C-RAN advantages: limiting multiplexing gain of BBU resources, limited enhancing of collaborative techniques.

- *Constant bit rate, independent on user traffic.* CPRI line bit rate is constant, independent on whether there is any user activity within a cell. A solution scaling the bit rate depending on user activity is desired.

- *One-to-one mapping RRH-BBU.* With CPRI each RRH is assigned to one BBU. For load balancing it is desired to be able to move RRH affiliation between pools.

- *FH bandwidth scales with cell configuration.* Current CPRI line rate depend on cell configuration: carrier bandwidth, number of sectors and number of antennas. Already for LTE-A up to 100 MHz can be used, and for Massive MIMO 100+ antennas can be expected. That yields 100+Gbps of data per cell, which is not feasible. Solution depending on user activity, and operating on lower bit rates is desired. Compression techniques presented in Section 2.4.4 received a lot of attention, however, none of the solutions are economically viable or flexible to meet 5G requirements. Therefore more disruptive solutions are needed.

Before 2013, to address the challenge of optimizing the fronthaul bit rate and flexibility various compression techniques were considered. As more disruptive methods were needed to achieve higher data rate reduction, nowadays (from 2013 onwards) a new split between RRH and BBU functionality is under extensive analysis by Alcatel Lucent Bell Labs [146], [147], NGMN [68], Small Cell Forum [148], [149] and many others. A new working group – NGFI – is under preparation (2015) under the sponsorship of an IEEE-SA Standards Sponsor with founding members of AT&T, Huawei, CMCC and Broadcom (more to join) [46]. Its goal is to encourage discussion on optimal functional splits between the pool and the cell site to address the above mentioned shortcomings of traditional fronthaul. The considered functional splits are marked with arrows in Figure 2.12. Functions to the left of each arrow will be executed centrally, while functions to the right of the arrow will be executed by the device at the cell site. Split can be done per cell [148], or even per bearer [150]. It may also be beneficial to implement different split for DL and UL. Most likely a few functional splits will be implemented per BBU pool, as one solution does not fit all the deployment scenarios. That requires a variety of devices at the cell site that will have the remaining functions implemented. With the current functional split (BB-RF), RRH could be standard-agnostic, especially when high frequency range was supported. Moving parts of L1 (and higher layers) to the devices at the cell site makes them dependent on mobile network standards.

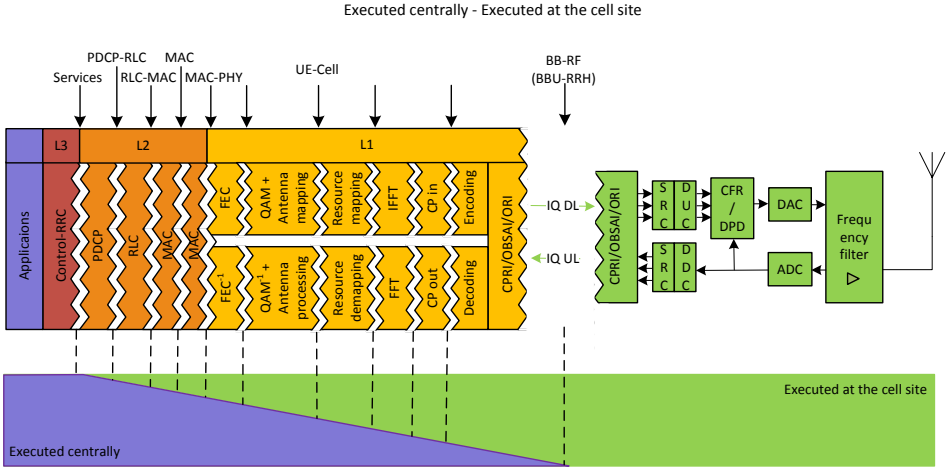


Figure 2.12: Possible functional splits for C-RAN

Various functional splits pose different throughput and delay requirements, as presented in Table 2.7. To benefit the most from centralization, the lowest split is rec-

ommended, i.e. the closest to the current BB-RF split. However, in order to save on bandwidth higher splits can be more applicable. User Equipment (UE)-Cell split (separating user and cell specific functionalities) is the lowest one for which data will have a variable bit rate, dependent on user traffic. Moreover, higher splits allow for higher fronthaul latency, e.g., 30 ms for PDCP-RLC split.

Table 2.7: Requirements for different functional splits [148] for the LTE protocol stack

Split	Latency	DL bandwidth	UL bandwidth
PDCP-RLC	30 ms	151 Mbps	48 Mbps
MAC	6 ms	151 Mbps	49 Mbps
MAC-PHY	2 ms	152 Mbps	49 Mbps
UE-Cell	250 μs	1075 Mbps	922 Mbps
BB-RF	250 μs	2457.6 Mbps	2457.6 Mbps

2.10.2 Ethernet-based fronthaul

With the new functional split, variable bit rate data streams on the fronthaul network are expected. The transport will be packet-based to optimized resource utilization in the network. Existing Ethernet networks can be reused for C-RAN fronthaul to leverage Ethernet's scalability and cost advantages. More on Ethernet-based fronthaul can be found in Chapter 4.

2.10.3 Centralization and decentralization trend

Looking at the base station architecture evolution, trends on centralization and decentralization can be observed as presented in Figures 2.13, 2.14. For an UMTS network, an example of a 3G mobile network, base station (NodeB) is separated from the network controller (Radio Network Controller (RNC)). For an LTE network those functionalities are distributed to each cell, and co-located in one logical node - eNodeB - in order to speed up scheduling and associated processing. As those functions are executed locally in eNodeB it shows an example of decentralization.

The physical architecture of a base station of 2-4G can be either a traditional base station (BB and RF units co-located), base station with RRH (decentralized) or C-RAN (decentralized looking at separation between RRH and BBU, but centralized BBUs). Towards 5G networks the trend is to move a part of BBU pool functionalities to the cell site, thereby resulting in the decentralization of a part of the pool. The actual point of functional split is moving over the years towards finding the optimal one for technologies and capacities available at given times, for given deployments.

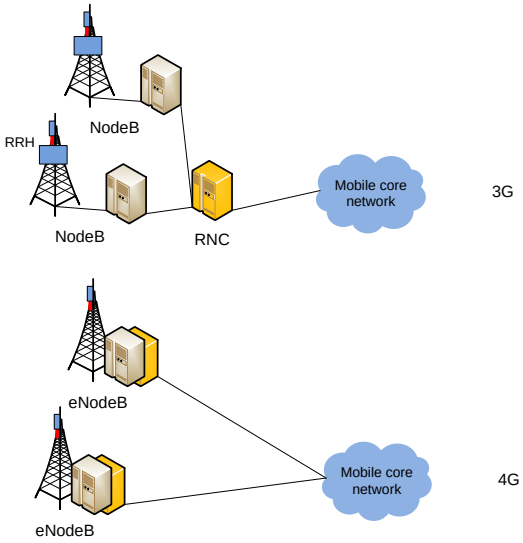


Figure 2.13: Decentralization of logical functions

2.11 Summary

This chapter presents a detailed overview of the novel mobile network architecture called C-RAN and discusses the advantages and challenges that need to be solved before its benefits can be fully exploited. C-RAN has the potential to reduce the network deployment and operation cost and, at the same time, improve system, mobility and coverage performance as well as energy efficiency.

The work towards resolving C-RAN challenges has been presented. Critical aspects such as the need for increased capacity in the fronthaul, virtualization techniques for the BBU pool and hardware implementation have been discussed. First prototypes and field trials of networks based on C-RAN have also been presented, together with most likely deployment scenarios.

While the concept of C-RAN has been clearly defined, more research is needed to find an optimal architecture that maximizes the benefits behind C-RAN. Mobile network operators as well as the telecommunication industry show a very high interest in C-RAN due to the fact that it offers potential cost savings, improved network performance and possibility to offer IaaS. However, the implementation of C-RAN needs to be justified by particular network operators taking into account available fronthaul network capacity as well as cost of virtualization of BBU resources. As the required fronthaul capacity is one of the main challenges for C-RAN deployments, the work on defining the new functional split is of utter importance.

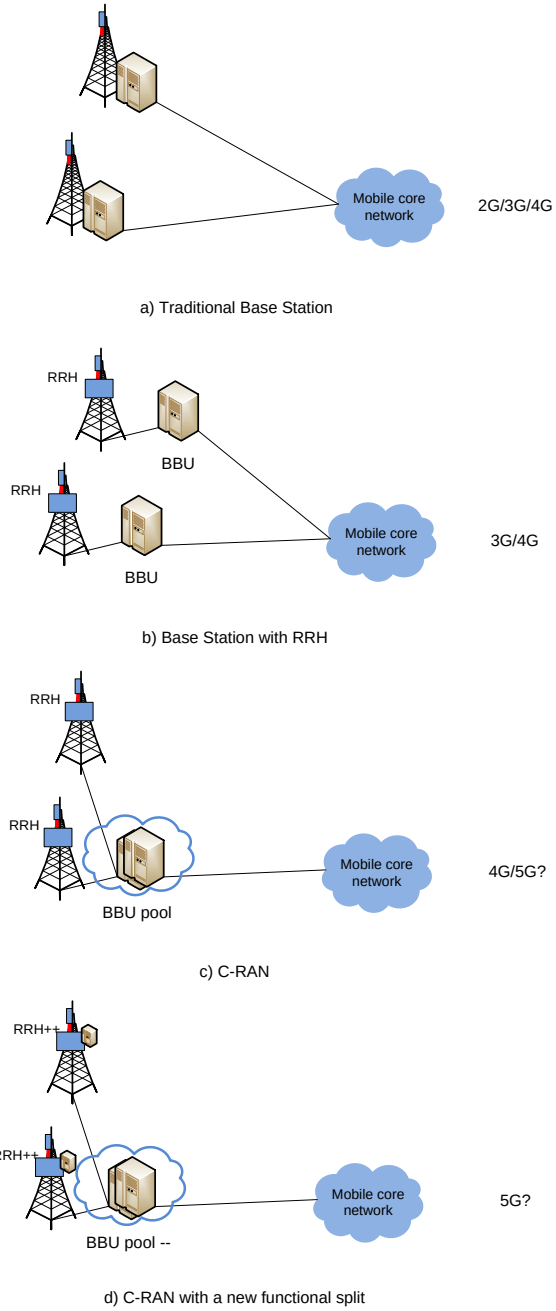


Figure 2.14: Decentralization, centralization and further decentralization of physical deployments, popular for given generations of mobile networks

CHAPTER 3

Multiplexing gains in Cloud RAN

Multiplexing gains can be achieved everywhere where resources are shared - aggregated by an aggregation point - and not occupied 100% of the time. This applies to printers shared by workers in one building, Internet packet-based networks where single lines are aggregated by e.g., a switch or a router as well as in C-RAN where BBUs are aggregated in a pool or when variable bit rate fronthaul traffic streams are aggregated.

By allowing many users and base stations to share network resources, a multiplexing gain can be achieved, as they will request peak data rates at different times. The multiplexing gain comes from traffic independence and from 1) burstiness of the traffic, and 2) the tidal effect - daily traffic variations between office and residential cells [20], described broader in Section 2.2.1. Given the fact that cells from metropolitan area can be connected to one BBU pool (maximum distance between RRH and BBU is required to be below 20 km) it is realistic to account for office and residential cells to be included. The tidal effect is one of the main motivations for introducing C-RAN [20].

The art of network/BBU dimensioning is to provide a network that is cost-effective for the operator and at the same time provides a reasonable QoS for users. As for any other shared resources, multiplexing enables to serve the same amount of users with less equipment. Multiplexing gain indicates the savings that come from the less equipment required to serve the same number of users. In this way the cost of deploying BBU pools and fronthaul links (CAPEX) will be lower. That will lead to energy savings, as fewer BBU units and fronthaul links need to be supplied with electricity (OPEX).

Multiplexing gain, along with enhancing cooperation techniques, co-locating BBUs, using RRHs as well as enabling easier deployment and maintenance are the major advantages of C-RAN. However, operators need to evaluate those advantages against the costs of introducing and maintaining C-RAN. This chapter presents several studies on quantifying the multiplexing gain that can serve as an input to the overall equation evaluating benefits and costs of C-RAN. First, multiplexing gains for different functional splits are presented in Section 3.2.3. Second, the methodology used to evaluate multiplexing gains is presented in Section 3.1. Section 3.2 presents two initial evaluations of statistical multiplexing gain of BBUs in C-RAN using 1) simple calculations, based on daily traffic load distribution, and 2) a simple network model, with real application definitions, however, without protocol processing. Section 3.3 presents results that include protocol processing and explore the tidal effect. Section 3.4 present results exploring different application mixes and presents additional conclusions on the impact of network dimensioning on users' QoS. Section 3.5 provides a discussion on the results, giving a

comprehensive analysis of the results obtained in various projects with different models used and exposing traffic burstiness and tidal effect. Section 3.7 summarizes this chapter.

A part of this section, Sections 3.2.3 - 3.1 and 3.4 were previously published in [13]. Section 3.2.2 was previously published in [11], while Section 3.3 was previously published in [12]. All of them are tailored and updated to fit this chapter.

3.1 Methodology

This chapter addresses the topic of multiplexing opportunities that are possible in C-RAN inside a BBU pool and on a fronthaul network. The conclusions are drawn based on the observation of the traffic aggregation properties in the simulated models and are an approximation of actual multiplexing gains on BBUs and fronthaul. This section aims at: 1) defining terms: multiplexing gain and pooling gain, 2) detailing elements of the LTE protocol stack and their impact on traffic shape, 3) explaining the relationship between the traffic aggregation properties and the gain in power consumption, processing resource and transport.

3.1.1 Definitions

This section defines Multiplexing gain (MG) and Pooling gain (PG), both in terms of processing resources and power.

MG - ratio between the sum of single resources and aggregated resources, showing how many times fewer resources are needed to satisfy users' needs when they are aggregated. Equation 3.1 shows an example for fronthaul resources, e.g. throughput. Link resources specify sufficient bandwidth for a given deployment. They can be defined in several ways, as providing peak throughput requested by users is costly: 1) the 95th percentile of requested throughput (used in Section 3.4.4), 2) mean plus standard deviation of requested resources (used in the teletraffic approach in [13]), 3) peak requested throughput averaged over given time (used in Sections 3.3, 3.2 and 3.4.2), and 4) link data rate for which the application layer delay is acceptable (used in Section 3.4.3). MG can refer to any resources, however, in the thesis it refers to savings on the transport resources. They are possible when several links are aggregated.

$$MG = \frac{\sum^{cells} SingleCellLinkResources}{AggregatedLinkResources} \tag{3.1}$$

PG - savings on BBUs comparing C-RAN with RAN (base station architecture with or without RRH, but not centralized). We can distinguish pooling gains in terms of processing, computational resources – $PG_{processing}$ and in terms of power savings – PG_{power}. Definition of $PG_{processing}$ is similar as for multiplexing gain but instead of Link Resources, BBU resources are considered, as presented in Equation 3.2. PG_{power} is defined in 3.3 in terms of power consumed by both architectures.

$$PG_{processing} = \frac{\sum^{cells} BBResources_{RAN}}{BBResources_{BBUpool}} \tag{3.2}$$

$$PG_{power} = \frac{\sum^{cells} BBPower_{RAN}}{BBPower_{C-RAN}} \tag{3.3}$$

The multiplexing gain value obtained in the thesis refers to MG. Section 3.1.3 describes the relation between MG, $PG_{processing}$ and PG_{power}.

3.1.2 Processing stack in LTE

This section describes the LTE base station processing stack, listing how each layer impacts traffic shape and thereby possible pooling gains. The detailed analysis on the bit rate after each layer processing has been provided by [147], [148] and [46].

LTE base station is responsible for terminating S1 interface on network side and Uu interface towards the user. Towards Uu interface, on the data processing plane, several steps need to be completed to prepare the data to be sent on an air interface. Figure 3.1 illustrates processing on L2 and L1.

The following steps need to be completed. Those steps can be divided into two groups: user-processing functionalities - above antenna mapping (inclusive), and, cell-processing functionalities - below resource mapping (inclusive) [148], [147], [46].

- Packet Data Convergence Protocol (PDCP) - it is responsible for integrity protection and verification of control plane data, (de-) ciphering of user and control plane data, it performs Internet Protocol (IP) header (de-) compression using RObust Header Compression (ROHC) protocol, it assures in-sequence delivery of upper layer Protocol Data Units Packet Data Units (PDUs) and duplicate elimination of lower layer Service Data Units Service Data Units (SDUs) [152]. Variable bit rate traffic is expected, following the users' activity, therefore it is assumed that the PG on this layer is equal to MG.

- Radio Link Control (RLC) - it is responsible for reliable and in-sequence transfer of information, eliminating duplicates and supporting segmentation and concatenation. It provides data transfer in three modes: Transparent Mode (TM), Unacknowledged Mode (UM) and Acknowledged Mode (AM). For TM it just transfers the upper layer PDUs. For UM it performs concatenation, segmentation and reassembly of RLC SDUs, reordering of RLC PDUs, duplicate detection and RLC SDU discard. On top of that, for AM error correction is done through Automatic Repeat-reQuest (ARQ), re-segmentation of RLC PDUs and protocol error detection [153]. Variable bit rate traffic is expected, following the users' activity. The above mentioned processing will change data pattern comparing to the traffic present at the PDCP

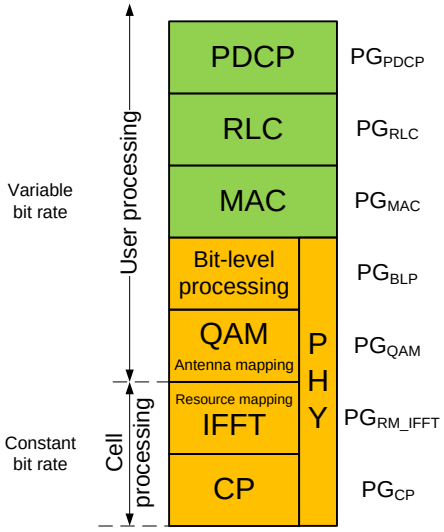


Figure 3.1: Layer 2 (green) and Layer 1 (yellow) of user-plane processing in DL in a LTE base station towards air interface. Possible PGs are indicated. Based on [151], [148], [147], [46].

layer. Still, it is assumed that the PG on this layer is equal to MG due to variable bit rate traffic.

- Media Access Control (MAC) - this layer is responsible for data transfer and radio resource allocation. It maps logical channels and transport channels into which information is organized, as well as multiplexes different logical channels into transport blocks to be delivered to the physical channel on transport channels, and demultiplexes information from different logical channels from transport blocks delivered from the physical layer on transport channels. It reports scheduling information, does error correction by Hybrid ARQ (HARQ), performs priority handling between UEs and between logical channels of one UE, and it selects the transport format [154]. Variable bit rate traffic is expected, following the users' activity. Above mentioned processing will change data pattern comparing to the traffic present at the RLC layer. Still, it is assumed that the PG on this layer is equal to MG due to variable bit rate traffic.

- various bit level processing takes place including Cyclic Redundancy Check (CRC) insertion, channel coding, HARQ and scrambling, increasing the traffic volume [151]. Still, a variable bit rate traffic is expected, following the users' activity. This functionalities can be put under the umbrella of Forward Error Correction (FEC). Above mentioned processing will change data pattern comparing to the traffic present at the MAC layer. Still, it is assumed that the PG on this layer is equal to MG due to variable bit rate traffic.

- Quadrature Amplitude Modulation (QAM) - downlink data modulation transforms a block of scrambled bits to a corresponding block of complex modulation symbols [151]. A variable bit rate traffic is expected in a form of IQ data, therefore it is assumed that the PG on this layer is equal to MG.

- the antenna mapping processes the modulation symbols corresponding to one or two transport blocks and maps the result to different antenna ports [151]. A variable bit rate traffic is expected in a form of IQ data, therefore it is assumed that the PG on this layer is equal to MG.

- the resource-block mapping takes the symbols to be transmitted on each antenna port and maps them to the resource elements of the set of resource blocks assigned by the MAC scheduler for the transmission. In order to generate an Orthogonal Frequency-Division Multiple (OFDM) signal IFFT is performed [151] resulting in constant bit rate stream of IQ data, therefore it is assumed that the PG on this layer is equal to one.

- cyclic-prefix (CP) insertion - the last part of the OFDM symbol is copied and inserted at the beginning of the OFDM symbol [151]. The data remains constant bit rate, therefore it is assumed that the PG on this layer is equal to one.

For each of these steps computational resources are required. The BBU pool needs to be dimensioned according to the planned user activity in order to meet strict real-time requirements of LTE processing. Depending on the functional split, for each of the layers, when it is included in the BBU pool, pooling gain is possible, both in terms of processing resources they require and power those resources consume: PG_{PDCP}, PG_{RLC}, PG_{MAC}, PG_{BLP}, PG_{QAM}, PG_{RM_IFFT} and PG_{CP}. Section 3.1.3 elaborates on such gains, splitting functionalities in user-processing and cell-processing. Assumed pooling gain on user-processing resources equals to MG while pooling gain on cell-processing resources will be equal to one, as listed in Table 3.1. In fact PG_{PDCP} will be closest to MG as the data is closest to the backhaul data. Processing by lower layers will change traffic properties, therefore pooling gain can be approximated to MG.

3.1.3 Gains in C-RAN

From Section 3.1.2 it can be seen which functionalities process variable and which constant bit rate traffic. This has an impact on whether pooling gains can be obtained.

Table 3.1: Assumed pooling gains on different layers of the LTE protocol stack.

PG	Value
PG_{PDCP}	MG
PG_{RLC}	MG
PG_{MAC}	MG
PG_{BLP}	MG
PG_{QAM}	MG
PG_{RM_IFFT}	1
PG_{CP}	1

Subsections below explain sources and nature of gains in processing resources, power consumption and transport network resources. The relation between MG, $PG_{processing}$ and PG_{power} is described.

3.1.3.1 Gains in processing resources

Those gains are referred to as pooling gains.

The BBU pool needs to be dimensioned according to the planned user activity. The more traffic is expected, the more computing resources will need to be available (e.g. rented from a cloud computing provider). Modules for control plane BB_{ctrl} will need to be provided, for cell-processing – traffic-independent functionalities BB_c, like FFT/IFFT, decoding/encoding as well as for user-processing – traffic dependent resources BB_u, like Bit-level processing (BLP). Equation 3.2 can be further expanded into equation 3.4 by inserting these components. Control resources, if dimensioned precisely, will be similar in RAN and C-RAN. In fact, in case of C-RAN, more control information may be needed to coordinate cells for e.g. CoMP. Cell-processing resources, also if dimensioned precisely will be similar in RAN and C-RAN. However, user-processing resources could be reduced by the factor of MG as listed in equation 3.5.

$$PG_{processing} = \frac{\sum^{cells} BBResources_{RAN}}{BBResources_{BBUpool}} =$$

$$= \frac{\sum^{cells} (BB_{RAN_{ctrl}} + BB_{RAN_c} + BB_{RAN_u})}{BB_{BBUpool_{ctrl}} + BB_{BBUpool_c} + BB_{BBUpool_u}} \approx \tag{3.4}$$

$$\approx \frac{\sum^{cells} BB_{RAN_{ctrl}} + \sum^{cells} BB_{RAN_c} + \sum^{cells} BB_{RAN_u}}{\sum^{cells} BB_{RAN_{ctrl}} + \sum^{cells} BB_{RAN_c} + \frac{\sum^{cells} BB_{RAN_u}}{MG}} \tag{3.5}$$

Therefore it can be seen that multiplexing gain calculated in the PhD project will affect only the amount of processing resources required for user-processing modules, leaving the fixed part aside. The factor of MG cannot be applied directly, as real life implementation needs to respond in real time to changing traffic, while averaged values were evaluated in the project, hence the almost equal to sign between equations 3.4 and 3.5. However, referring to the equation 3.2, it is worth noticing that, in practice, BBUs will always be dimensioned with a margin comparing to planned consumption, allowing to accommodate higher traffic peaks if they occur and to account for forecasted overall traffic growth. When BBUs are aggregated in a pool, such a margin could be shared, allowing additional pooling gain in C-RAN comparing to RAN. Such an additional margin is especially applicable to resources required to support traffic peaks in different cells. In other words, capacity can be scaled based on average utilization across all cells, rather than all cells peak utilization. Moreover, the processing power can be dynamically shifted to heavily loaded cells as the need arises.

In [41], Werthmann *et al.* show results on pooling gain, analyzing compute resource utilization, in Giga Operations Per Second (GOPS), for different number of sectors aggregated into a BBU pool. In [155], Desset *et al.* elaborate on how many GOPS are needed for different base station modules, including:

- Digital-to-Analog Converter (DPD)
- Filter: up/down-sampling and filtering
- CRPI/SERDES: serial link to backbone network
- OFDM: FFT and OFDM-specific processing
- FD: Frequency-Domain processing (mapping/demapping, MIMO equalization); it is split into two parts, scaling linearly and non-linearly with the number of antennas

Table 3.2: Estimations of baseband complexity in GOPS of cell- and user-processing for UL and DL and different cell sizes. Numbers are taken from [155].

	DL		UL	
	macro	femto	macro	femto
$GOPS_c$	830	180	700	240
$GOPS_u$	30	25	140	120
$\%GOPS_u$	3%	12%	17%	33%

- FEC

- CPU: platform control processor.

For the reference case, assuming 20 MHz bandwidth, single-antenna, 64-QAM, rate-1 encoding and a load of 100%, and treating DPD, filtering, SERDES, OFDM, and linear part of FD as cell-processing, and non-linear part of FD and FEC as user-processing, Table 3.2 summarizes estimations on GOPS for UL and DL for macro and femto cell for cell-processing $GOPS_c$ and for user-processing $GOPS_u$. It would require further investigation on how to split platform control processor resources. It can be concluded that especially for downlink, share of user-processing resources is small comparing to cell-processing resources (3-12%). For uplink processing relatively more can be saved on user-processing (17-33%).

To conclude, MG calculated in the project applies only to user-processing resources, which are a fraction of total resources. Moreover, as mentioned above there are factors implying that MG calculated in the project is actually both lower (protocol processing) and higher (dimensioning margin) in reality. Given the complexity of the system and amount of factors influencing the final result, obtained MG gives only a contribution to the approximation of the real-life $PG_{processing}$.

3.1.3.2 Gains in power consumption

A base station consumes power on the following:

- powering the main equipment (51% of the cell site power consumption for traditional base station in China Mobile [20], 83% for Vodafone [156] found in [157] and [158])

- air conditioning (46% of the cell site power consumption for traditional base station in China Mobile [20], 10-25% for Vodafone [156])

- fronthaul network transmission (in case of C-RAN).

Power consumption of the main equipment can be further broken down into:

- $Power_{RRH}$ powering RRH, necessary to assure coverage, not dependent on the traffic load, includes RF processing and Power Amplifier (PA);

- $Power_{ctrl}$ powering modules responsible for control-plane processing

- $Power_c$ powering modules responsible for baseband processing of cell-processing functionalities, not dependent on the traffic load, including resource mapping, IFFT and CP

- $Power_u$ powering modules responsible for baseband processing of user-processing functionalities, dependent on the traffic load, including PDCP, RLC, MAC, BLP, QAM and antenna mapping.

According to [156] main equipment consumes power on:

- PA - 50-80% of total base station consumption

- signal processing - 5-15% of total base station consumption

- power supply - 5-10% of total base station consumption

. Based on this data $Power_{ctrl} + Power_c + Power_u$ account for 5-10% of total base station power consumption. In [159] Auer *et al.* reports that 2-24 % of total base station consumption are spent on baseband processing.

Similar as for processing resources, equation 3.3 can be further expanded into equation 3.6 by inserting these components. Also, only user-processing resources could be reduced by the factor of MG as listed in equation 3.7.

$$PG_{power} = \frac{\sum^{cells} BBPower_{RAN}}{BBPower_{C-RAN}} =$$

$$= \frac{\sum^{cells} (Power_{RAN_{ctrl}} + Power_{RAN_c} + Power_{RAN_u})}{Power_{BBUpool_{ctrl}} + Power_{BBUpool_c} + Power_{BBUpool_u}} \approx \quad (3.6)$$

$$\approx \frac{\sum^{cells} Power_{RAN_{ctrl}} + \sum^{cells} Power_{RAN_c} + \sum^{cells} Power_{RAN_u}}{\sum^{cells} Power_{RAN_{ctrl}} + \sum^{cells} Power_{RAN_c} + \frac{\sum^{cells} Power_{RAN_u}}{MG}} \quad (3.7)$$

In [155], Desset *et al.* propose a more detailed power model of a base station, listing factors that impact power consumption. Power consumption is proportional to GOPS: the number of operations that can be performed per second and per Watt is 40

GOPS/Ws for large base stations and default technology, here i.e., 65 nm General Purpose Complementary metal-oxide-semiconductor (CMOS), and it is three times larger for pico and femto cells. Therefore PG_{power} on user-processing resources can be reduced by the same factor as $PG_{processing}$ on user-processing resources.

To conclude, PG_{power}, similarly to $PG_{processing}$, accounting for 2-24% of base station power consumption, is affected by MG by reducing power spent by user-processing resources. Model including the whole LTE processing stack would be needed to determine overall savings more precisely, as well as detailed information on how user traffic impacts power consumption.

3.1.3.3 Gains in transport

The fronthaul network capacity needs to be dimensioned according to the planned user activity. In a packet-based fronthaul, traffic to different cells can be aggregated. For UE-Cell functional split traffic will be variable bit rate and will resemble traffic from the simulations, however simplifications mentioned in Section 3.5.6 apply. Aggregated link capacity can be directly reduced by MG times, if the QoS is assured.

3.1.4 Approach

Mathematical, teletraffic theories have been used to calculate an overbooking factor [160] that dimensions the link, based on the definition of an effective bandwidth [161]. They provide an important indication when the fundamentals of the networks are studied.

However, teletraffic theories focus on well-defined traffic models, such as ON - OFF source traffic e.g., Interrupted Poisson Process, Interrupted Bernoulli Process [161]. As such, they do not capture all the aspects of real-life networks. In current and future mobile networks there is a big variety of applications and the traffic varies throughout the day depending on cell location (office or residential). In order to capture this heterogeneity the analysis is done in a discrete event-based simulator - OPNET [162]. Such scenario with detailed and heterogeneous traffic definition is especially important to evaluate the UE-Cell split (introduced in Figure 2.12. Simulations allow to capture protocol interactions and thereby observe implications of different network architectures on end-to-end delay seen at the application layer. On the other hand, a mathematical approach allows to create simpler models that can run with lower simulation time, thereby enabling testing more extended scenarios, e.g., with more cells. Both approaches are important to compare and validate the results. This chapter reports simulation results that have been validated analytically by collaborators, as presented in Section 3.3.2.

Quantifying multiplexing gains has been addressed by research and industry communities. Related work is presented in Section 2.2.1 and 3.4.4.

3.2 Origins of multiplexing gain

Sections 3.2.1, 3.2.2 and 3.2.3 present different sources of multiplexing gain: tidal effect, traffic burstiness and C-RAN functional splits. Section 3.2.1 explores average daily loads in office, residential and commercial areas via simple calculations. Section 3.2.2 presents results of a simple network model where applications have been modeled on top of daily traffic loads. Section 3.2.3 presents theoretical considerations.

3.2.1 Tidal effect

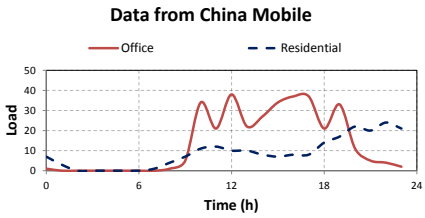
A tidal effect is one of the main factors affecting multiplexing gain. People tend to be in at least two places during the day (e.g., home and work), therefore as a rule of thumb they could use 1 BBU in C-RAN instead of 2 different BBUs in RAN, what can be counted as a multiplexing gain of 2, based on equation 3.2. However, they could use the network differently in different places causing different amount and characteristics of traffic. Therefore the value of multiplexing gain could be slightly lower or higher than 2.

Two examples of data sets have been used to calculate multiplexing gain according to equations 3.2, where $BBResources$ correspond to traffic load:

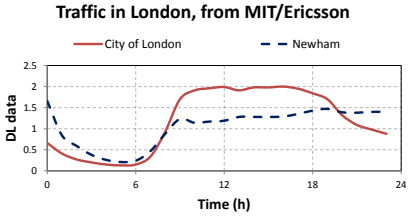
- from China Mobile, presented in Section 2.2.1
- from Ericsson and MIT analyzed within the project "A tale of many cities" [163]. Data shows daily, weekly, monthly and overall traffic patterns in terms of calls, SMS, DL data, UL data and requests from various districts of London, New York, Los Angeles and Hong Kong. DL data was analyzed.

Sample values are shown in Figure 3.2.

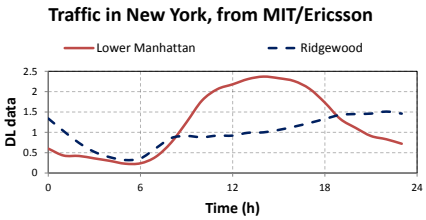
Figure 3.3 shows multiplexing gain calculated based on data from China Mobile and Ericsson/MIT (London, New York and Hong Kong) for varying percentage of office vs residential cells. Traffic from Ericsson/MIT includes DL data for a typical week in two different office/residential areas: City of London and Newham for London, Battery Park City - Tribeca (Lower Manhattan) and Ridgewood for New York, as well as "Central and Western" and Yuen Long for Hong Kong. Tuesday is taken into account to show a typical day in the middle of the week. For 100% of residential (0% office) and 100% of office cells multiplexing gain equals to one, which is expected, since similar daily traffic distributions are aggregated. In other cases values are within 1.0 - 1.32 with peak values for 20-30% of office cells for all the cities except for Hong Kong, where peak occurs for 60% of office cells. The differences are due to different daily traffic distribution in different cities. Figure 3.2 shows an example. Values of traffic load cannot be directly compared as they come without absolute values, but the trend lines can. In data from China Mobile, presented in Figure 3.2(a), traffic during business hours in residential areas is significantly lower than in the office areas and the traffic load is complimentary in office and residential areas through the day, hence a higher multiplexing gain value. On the contrary, in London, Figure 3.2(b), traffic is more uniformly distributed between office



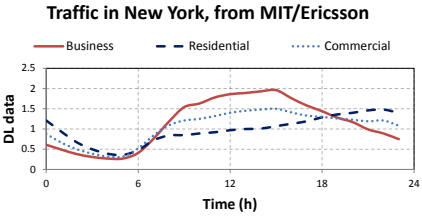
(a) Overall traffic reported by China Mobile



(b) Example of traffic in an office and residential cell in London



(c) Example of traffic in office and residential cell in New York



(d) Overall traffic in New York

Figure 3.2: Daily traffic distributions.

and residential areas, only peaking in the office area in business hours, hence lower values of multiplexing gain. Data for Los Angeles: Downtown and Florence, resembles data from London. Values for New York, Figure 3.2(c) lie in between values for London and data from China Mobile, as the curves of traffic.

Typically in the cities there are not only residential and office areas but also commercial areas, like shopping malls, movie theaters as well as parks and mixed areas. They will all have different daily and weekly traffic distributions that will affect the value of multiplexing gain. Based on the available data impact of commercial sites is estimated as shown in Figure 3.4. It is based on the overall data for New York from MIT/Ericsson presented in Figure 3.2(d). Analyzed deployments consist of office, residential and commercial cells. Values vary between 1.0, for 100% of one cell type, to 1.21 for 30% of office cells, 70% residential and 0% commercial. As shown in Figure 3.2(d) commercial areas in New York have a uniform traffic distribution in the daytime therefore they lower the multiplexing gain.

To conclude, in order to maximize multiplexing gains it is best to combine in a BBU pool cells whose daily traffic distributions are complementary i.e. traffic is low in some

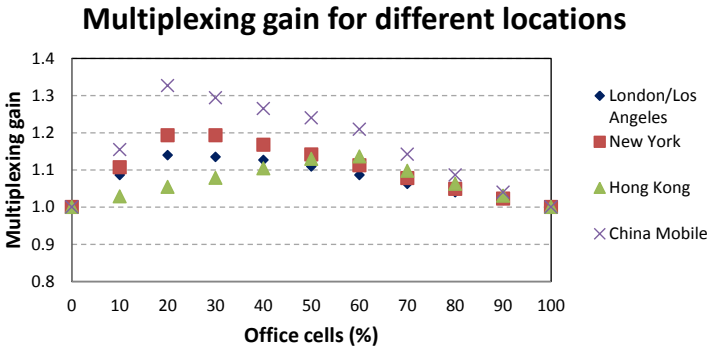


Figure 3.3: Multiplexing gains for different locations based on daily traffic distributions between office and residential cells.

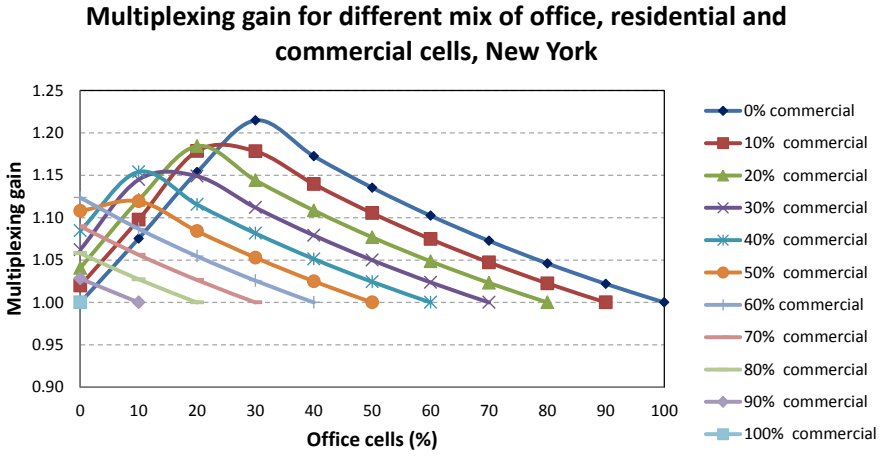


Figure 3.4: Multiplexing gains for different distributions between office, residential and commercial cells.

cells while it is high in others and vice versa.

3.2.2 Traffic burstiness

Multiplexing gain comes not only from the tidal effect, but also from traffic burstiness. Due to the fact that user activity varies in time, it is possible to allow many users to share the same link without too much queuing, in the same way as printers could be shared

by office workers on one floor. In order to capture traffic burstiness real-life application characteristics need to be taken into account.

Paper [11] presents an initial evaluation of the multiplexing gain including both the tidal effect and traffic burstiness. A real case scenario is modeled with the mobile traffic forecast for year 2017, a number of recommendations on traffic models, including a daily traffic variations between office and residential cells as presented in [20] and a proposed RAN and C-RAN implementation. RAN and C-RAN are modeled by nodes performing simple traffic aggregation. Multiplexing gain is evaluated as in equation 3.1, where $LinkResources$ are represented by the peak requested throughput.

The results show that the statistical multiplexing gain for user traffic in a C-RAN architecture, when traffic burstiness is taken into account, is $4.34, \sigma = 1.42$, compared to a traditional RAN architecture. Please refer to [11] for more details on the model and results.

3.2.3 Different functional splits

In a traditional base station or in a base station with RRH, for each cell, baseband processing resources are statically assigned to the RRH, as shown in Figure 3.5a. In C-RAN, presented in Figure 3.5d the baseband units are shared in a virtualized BBU pool, hence it is expected that in C-RAN the amount of processors needed to perform baseband processing will be lower compared to RAN. The CPRI protocol is constant bit rate, independent of user activity. Hence, there is no multiplexing gain on fronthaul links. This split is referred to as BB-RF as it separates baseband and radio frequency functionalities.

Several examples of functional splits are indicated in Figure 3.5: BB-RF (Figure 3.5d), discussed above, PDCP-RLC (Figure 3.5b) and UE-Cell (Figure 3.5c). With the *UE-Cell split* (separating user and cell specific functionalities) traffic between RRH and BBU is traffic-dependent, hence multiplexing gain can be expected both on BBU resources and also on the fronthaul links.

For *PDCP-RLC split*, the majority of data processing is executed at the cell site, only a small portion of it is done in the pool, hence a marginal BBU pool multiplexing gain. However, a variable bit rate traffic is transmitted on the fronthaul links, hence a possibility for a multiplexing gain on the fronthaul. This split leaves the MAC scheduling and PHY functionality to reside at the RRH, which limits the possibility of joint PHY processing and joint scheduling for multi-cell cooperation.

Generally, the more processing is centralized, the higher savings on BBU pool cost and benefits coming from centralization, but higher burden on fronthaul links. On the other hand, the more functionalities are left at the cell site, the lower savings on the BBU pool, but at the same time lower cost of fronthaul coming from multiplexing gain, lower bit rates and relaxed latency requirements, as indicated in Section 2.10.1.

The aggregated link in equation 3.1 represents fronthaul and BBU traffic for UE-Cell split, therefore the multiplexing gain on fronthaul links for UE-Cell split can be calculated

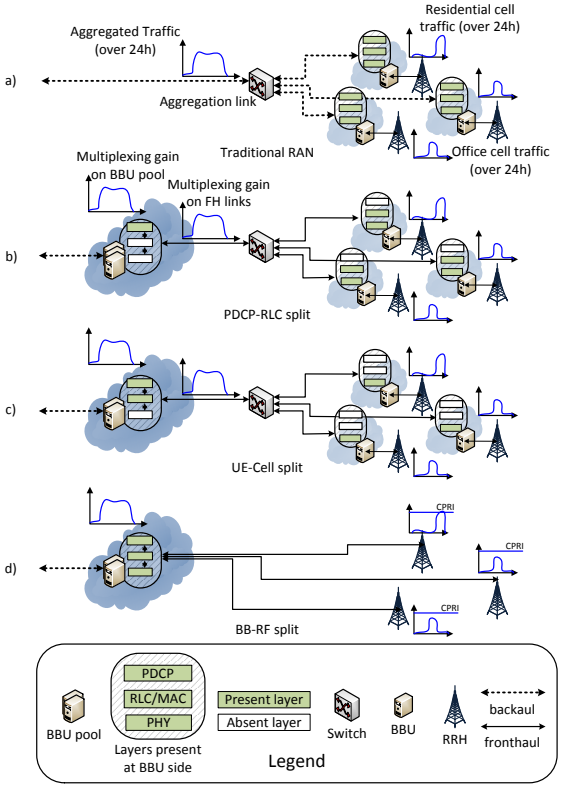


Figure 3.5: Possible multiplexing gains on BBU pool and fronthaul links depending on base station architecture.

$MG_{FH-UE-Cell}$. $MG_{FH-UE-Cell}$ comes straightforward from equation (3.1) where $LinkResources$ are quantified as throughput or data rate. Only traffic dependent resources are evaluated, therefore the comparison between single and aggregated link resources is analogical to comparing traffic on BBUs in RAN to BBU pool in C-RAN. As a consequence, multiplexing gain on fronthaul for UE-Cell split $MG_{FH-UE-Cell}$ is the same as multiplexing gain on BBU for UE-Cell split $MG_{BBU-UE-Cell}$ and BB-RF split $MG_{BBU-BB-RF}$. Later in this chapter, all these results will be referred to as MG, however the same conclusions appear to $MG_{FH-UE-Cell}$, $MG_{BBU-UE-Cell}$ and $MG_{BBU-BB-RF}$. Table 3.3 shows dependencies between the values.

Table 3.3: Multiplexing gains (MG) looking at traffic-dependent resources.

RAN Architecture	BBU	links
Traditional RAN	1 (no MG)	MG presented in this chapter apply
PDCP-RLC split	[13]	[13]
UE-Cell split	MG presented in this chapter apply	MG presented in this chapter apply
BB-RF	MG presented in this chapter apply	1 (no MG)

3.3 Exploring the tidal effect

As stated before, multiplexing gain comes from traffic burstiness and from the tidal effect. This section explores the tidal effect and summarizes efforts to find the most optimal mix of residential and office cells in order to maximize the multiplexing gain. The results can be applied to the BBU pool for BB-RF split.

3.3.1 Network model

Network simulations are carried out in OPNET Modeler inspired by a traffic forecasts for 2019 [164], [1]. Simulations are carried out for 16 hours to study the impact on varying traffic load throughout the day - from 8 a.m. to midnight. Traffic load for the average office and residential base station follows the trend observed in the network operated by China Mobile, presented in Figure 2.5 [20]. The parameters of the traffic models are summarized in Table 3.4. The actual applications run on top of the TCP/IP and Ethernet protocol stack, as shown in Figure 3.6. The simulations are ran for 10 cells, varying the number of office and residential base stations. Traffic is aggregated by the Ethernet switch. The peak throughput in downlink direction is observed on the links between the cells and the switch to obtain $SingleCellLinkResources$ and the peak throughput on the link between the Ethernet switch and the BBU pool to collect $SingleCellLinkResources$, in order to calculate multiplexing gain according to equation (3.1). Throughput measurements are taken every 576 s (100 measurement points in 16 hours). 24 runs with different seed values allow to obtain standard deviation of the results. Confidence intervals for 95% level are calculated using the Student's t distribution. Student's t distribution is used instead of a normal distribution because a relatively small number of samples (24) is analyzed.

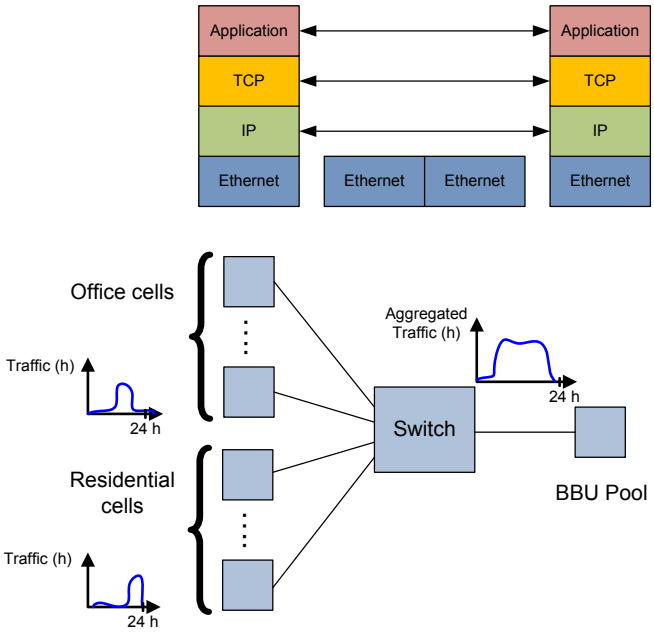


Figure 3.6: Network model used for simulations.

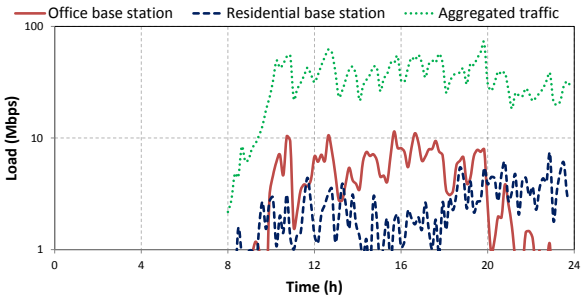


Figure 3.7: Modeled traffic from residential, office and aggregated traffic from 5 office and 5 residential cells.

Table 3.4: Traffic generation parameters for network modeling; C - Constant, E - Exponential, L - log-normal, U - uniform, UI - uniform integer

Traffic Parameters	Value	Distribution
Video application - 50% of total traffic		
Frame interarrival time	10 frames/sec	C
Frame size	128 x 120 pixels - 17280 B	C
Start time offset	min 50 s, max 3600 s	U
Duration of profile run	mean: 180 s	E
Inter-repetition time	mean: 1800 s	E
File sharing application - 40% of total traffic		
Inter request time	mean: 180 s [165]	P
File size mean	2 MB [165]	L
File size Standard Deviation	0.722 MB [165]	
Duration of profile run	mean: 1200 s	E
Inter-repetition time	300 s	E
Web browsing application - 10% of total traffic		
Page interarrival time	mean: 10 s	E
Page properties	main object: 1 KB	C
	Number of images: 7	C
	Size of images: min 2KB, max 10KB	UI

3.3.2 Discussion on the results exploring the tidal effect

Figure 3.7 shows modeled traffic from residential, office and aggregated traffic from 5 office and 5 residential cells. Please note the logarithmic scale.

Figure 3.8 summarizes the results on the statistical multiplexing gain for different percentage of office base stations for the studied traffic profile. It shows that the maximum statistical multiplexing gain is achieved for a BBU Pool serving 30% of office base stations and 70% of residential base stations. Multiplexing gain value of 1.6 corresponds to 38% of BBU savings. The way this and the following BBU savings are calculated is as presented in equation 3.8.

$$BBU_Saving = \frac{MG - 1}{MG} = \frac{1.6 - 1}{1.6} \tag{3.8}$$

Multiplexing gain for all the cases is above 1.2, corresponding to 17% of BBU savings. Different traffic profiles will affect the results, however the same model can be used to process different input data.

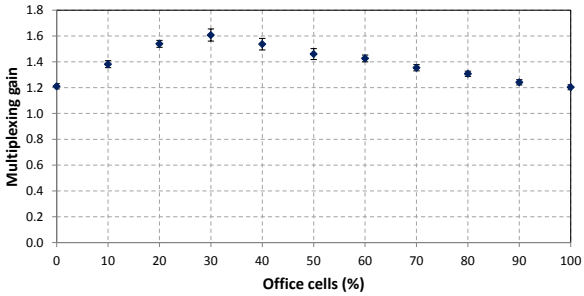


Figure 3.8: Optimal distribution of office and residential cells - simulation results. Confidence intervals for 95% level are shown.

Simulation results are verified and confirmed by collaborators, using 1) an analytical approach using calculations based on averaged throughput in [12], as well as, 2) using teletraffic methods in [45]. Looking at the statistical multiplexing gain that can be achieved in C-RAN under our assumptions, it reaches the maximum value for: 1) 30% of office cells concluding from the modeling approach, with the statistical multiplexing gain reaching 1.6, 2) 21% of office cells in the analytical approach resulting in 70% of BBU save, and 3) 22% of office cells in the teletraffic approach resulting in 84% of BBU saving. The analytical approach takes an average network load for each hour, teletraffic approach - aggregated traffic characteristics, while the modeling approach included burstiness resulting from application definitions and protocol stack processing. All three approaches show the same trend line for optimal office/residential cell mix peaking at 20-30% of office cells, which validates the results. Values of BBU saving are less straightforward to compare as they are obtained with different methods. The ones coming from multiplexing gain itself vary between 35-49%. Table 3.5 compares the results.

3.4 Exploring different resources measurement methods and application mixes

The previous section investigates multiplexing gains resulting from the tidal effect. In this section the most optimal office/residential cells mix is considered and the different methods of evaluating $LinkResources$ from equation 3.1 are used and compared. Moreover, the investigation focuses on the application mix impact, therefore the percentage of web and video traffic is varied, while the total offered traffic follows the daily load for each simulation run. Those results study the impact of traffic burstiness, therefore they can be applied to both BBU pool and fronthaul for UE-Cell split.

Table 3.5: BBU save for various office/residential cell mixes, measured using different methods.

Method	Maximum BBU save	% of office cells	Source
Discrete-event based simulations	38% coming from MG (MG=1.6)	30%	[12]
Analytical - averaged calculations	35% coming from MG (MG=1.54), 70% using authors method	21%	[12]
Teletraffic study based on a multi-dimensional loss system	49% coming from MG (MG=1.96), 84% using the Moe's principle	22%	[45]

3.4.1 Discrete event simulation model

A sample system is built with an architecture and protocol stack similar to the one used for our previous study [12]. Ten base stations are connected to one switch and then to a server, as presented in Figure 3.9. Traffic in the BBU pool and on the fronthaul links in UE-Cell split can be compared to Ethernet traffic on the MAC layer. PHY layer cell processing is done at the cell site leaving MAC-layer-like traffic on the fronthaul. Each base station is connected with a 1 Gbps link, as this could be a radio link throughput of LTE-A and initially data rate of the aggregated link is 10 Gbps not to create any bottleneck. There are three office and seven residential base stations, as this is the mix for which we observed the maximum multiplexing gain in our previous studies [12], [45]. Daily traffic load between office and residential cells varies according to [20], as in previous studies. We send video and web traffic according to the definitions presented in Table 3.6 to represent delay sensitive (ms level) and delay insensitive (s level) applications. Parameters for video traffic are based on [166] and for web traffic on [167] considering an average web page size growth between years 2007 and 2014 [168]. Values presented in the table represent traffic from 8 a.m. to 9 a.m. for office base stations (lowest load observed in the system) and are multiplied for other hours and residential base stations to reflect the daily load. Simulation parameters are presented in Table 3.7. No QoS-aware scheduling is done, the packets are processed in First Input First Output (FIFO) manner. This simple scheduling algorithm is used to show the emphasis on traffic aggregation, not scheduling as such.

3.4.2 Throughput measurements for quantifying multiplexing gains

Given the fact that the links between the base station and the aggregating switch have a data rate of 1 Gbps, it can be seen on ns level whether a bit is sent or not. LTE scheduling is done every 1 ms, therefore it is viable to measure not more often then once every

Table 3.6: Traffic generation parameters for network modeling; C - Constant, E - Exponential, L - log-normal, G - gamma, U - uniform

Traffic Parameters	Value, Distribution
Video application	
Frame interarrival time	10 frames/sec, C
Frame size	4193 B, C
Duration of video conference, 50% cases	mean: 16 s, variance: 5 s, N
Duration of video conference, 25% cases	mean: 208 s, variance: 3364 s, N
Duration of video conference, 25% cases	mean: 295 s, variance: 2500 s, N
Inter-repetition time	mean: 1250 s, E
Web browsing application	
Page interarrival time	mean: 28.5 s, variance: 1774966 s, L
Page properties	main object size mean: 63346 B, variance: 86205010504 B, max 6 MB, L
	Number of embedded objects scale: 40, shape: 0.1414, max 300, G
	embedded object size mean: 142103 B, variance: 2264446191523 B, max 6 MB, L

Table 3.7: Simulation parameters.

Parameter	Value
Modeler and simulation software	OPNET 17.5.A PL3
Simulated time	16 h
Seeds	24, random
Values per statistic	For throughput measurements every 10 ms, for delay measurements every 1 s

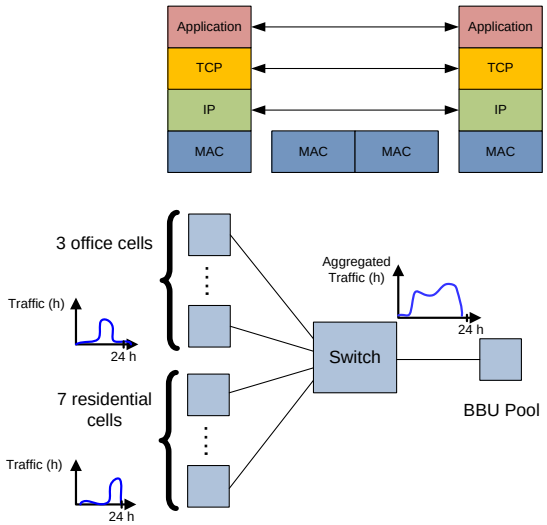


Figure 3.9: Network model used for simulations.

1 ms. For practical reasons, in order to be able to process the results efficiently, the data is collected every 10 ms. Operators will most likely not dimension their networks for peak user data measured over 1 ms, but allow some buffering, thereby saving the costs, although lowering the user data rate. Therefore different averaging is applied over simulated throughput as follows.

For each cell c and for the aggregated link a the data set resulting from simulations consists of $\frac{16hours}{10ms} = 5760000$ throughput measurements x measured at time t. An averaging window (bucket) is defined with width W such that for the samples (t_i, x_i) where $i = 0, 1, ..., n$ and $t_n - t_0 = W$. The averaging window size represents the networks ability to smoothen the traffic and has a similar function to a buffer. 16 hours simulated time is divided into such windows W and for each of them an average throughput is calculated $y = \frac{\sum_{i=0}^{n} x_i}{n}$. Out of all the y values, a maximum value of all the averages is found y_{max} for each cell $y_{max,c}$ and for an aggregated link $y_{max,a}$. Based on equation (3.1) MG_{AVG} is calculated as presented in equation (3.9).

$$MG_{AVG} = \frac{\sum_{c=1}^{cells} y_{max,c}}{y_{max,a}} \tag{3.9}$$

Values of MG_{AVG} coming from simulations for different web and video traffic

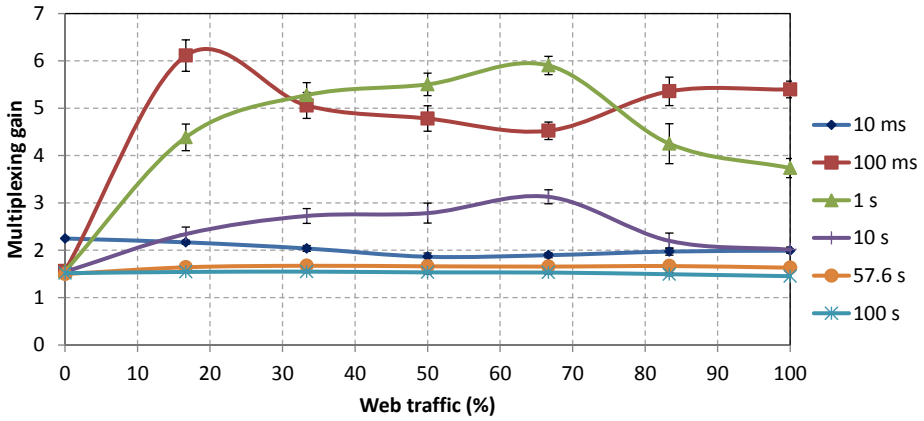


Figure 3.10: Multiplexing gain for for different percentage of web traffic in the system and different throughput averaging windows: $MG_{FH-UE-Cell}$ (10 ms, no averaging) and $MG_{FH-UE-Cell-AVG}$ (100 ms, 1 s, 10 s, 57 s and 100 s).

mixes are presented in Figure 3.10. Confidence intervals for 95% level are calculated using the Student's t distribution. Different series present data averaged over 100 ms, 1 s, 10 s, 57 s and 100 s (averaging window W). For 10 ms series throughput is not averaged, only the maximum values are taken for each cell and aggregated link to compute MG. Values vary for different mixes of web traffic. For no web traffic present in the network, the multiplexing gain has similar value within our averaging intervals, as video conferencing sessions have constant bit rates. As soon as web traffic is present (17-100%) multiplexing gain varies from 1.5 to 6 depending on the averaging window. It can be seen that multiplexing gain is very sensitive to the measurement interval. There is a clear dependence of the averaging period on the multiplexing gain.

In principle, if we take longer, up to infinite, averaging periods the multiplexing gain should be getting lower and reaching one, as the average bandwidth of an aggregation link will need to match the sum of average bandwidths of single links. Therefore it is not straightforward why the value is low for every 10 ms, then increases for 100 ms and 1 s and then lowers again. The possible cause could be that the Cumulative Distribution Function (CDF) of the throughput looks in a way that for 90% of the time the throughput to base stations is below 100 Mbps and aggregated throughput is below 1 Gbps, as presented in Figure 3.11. This indicates that by adequate dimensioning the multiplexing gain value can be different. Moreover, if the dimensioning is done according to the results from averaging over longer periods, the risk of dropped packets and connections will increase, as buffer sizes may be exceeded and packets may be dropped or users may not be satisfied

with the delay. In this study none of the packets were dropped. The averaging is done only in post processing of the data, so actually it is not verified what will be the impact of providing only the data rates as averaged on the application level delays. For video conferencing and web browsing averaging only up to 10 - 100 ms is safe, as application layer delays should not exceed the order of magnitude of 150 ms and 1 s, respectively. Delays on the application level will give an ultimate criterion for network dimensioning. The following section elaborates on them.

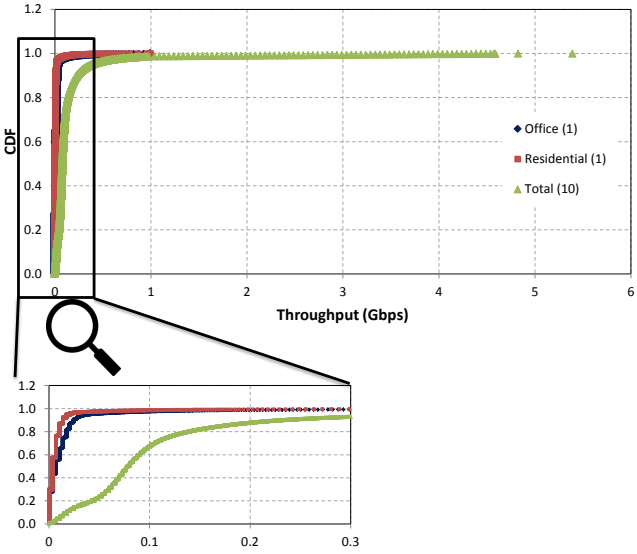


Figure 3.11: CDFs of throughput for an sample office and residential cell as well as total throughput for all ten cells for 50% web traffic mix.

3.4.3 Delay-based criteria for network dimensioning

The ultimate criterion to dimension the discussed links and BBU pool is to assure acceptable performance on the application level. For that web page response time and video packet end-to-end delay is checked via discrete event simulations. For 100-400 Mbps aggregated link data rate the delay differences are the highest and they reach the point when they become acceptable. The links are intentionally examined with fine granularity of throughputs (every 50 Mbps) as lines and computational power can be leased with fine granularity of data rates [169]. The results are presented in Figures 3.12 and 3.13. For web traffic 90th percentile of web page response time is below 1 s for link data rates $\geq$200

Mbps. For video conferencing 90th percentile of packet end-to-end delays are below 150 ms for link $\geq$200 Mbps.

As expected, the delays are lower when the offered link data rates are higher. The impact on delay is higher for the cases with less web traffic. It is due to the fact, that the more video traffic is present in the network, the delays are more sensitive to the aggregated link bandwidth. Small change of data rate affects delay considerably, even by a factor of 10 (for 17% of web traffic - 83% of video traffic). The reason could be that video occupies a link at a constant bit rate for at least a minute, so if the links are under-dimensioned queuing occurs. The conclusion is, that the more bursty the traffic is, the less sensitive is it to under-dimensioning. The more video traffic present in the network the dimensioning becomes more relevant for achieving quality of service. Traffic forecast [2] predicts that in 2020 60% of the mobile traffic will be video; it will, however, vary between light and heavy users.

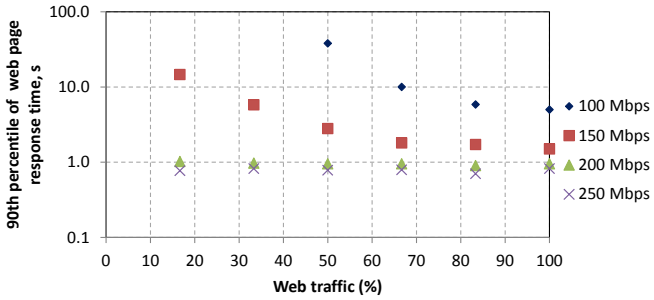


Figure 3.12: 90th percentile of web page response time for different percentage of web traffic in the system and for different aggregated link data rate.

3.4.4 Analysis based on the 95th percentile of requested throughput

The 80th, 90th, and 95th percentiles of the sums of single cells throughputs and of an aggregated link throughput are analyzed. The results are presented in Figure 3.14. The sum of the 80th and 90th percentiles are getting lower because the more web traffic is present in the network the lower the mean traffic, but the standard deviation gets higher. However, looking at the 80th and 90th percentiles on the aggregated link, it is getting higher because the peaks occur more often. The trend shown in Figures 3.12 and 3.13 is the same as for the sum of the 80th and 90th percentiles in Figure 3.14. The values closest to 200 Mbps proven to be sufficient from the delay analysis. They account for the higher one of those values: the 80th percentile on aggregated link or the sum of 90th percentiles. Therefore it can be concluded that in order to dimension fronthaul links and BBU, the

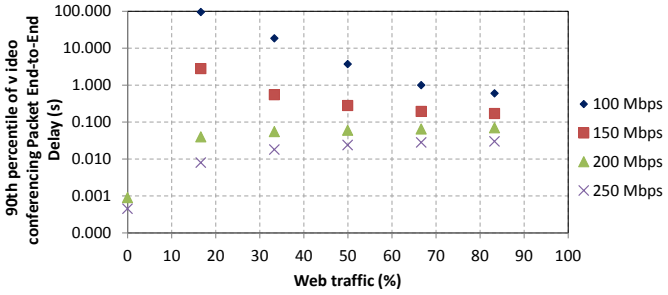


Figure 3.13: 90th percentile of video conferencing packet End-to-End delay for different percentage of web traffic in the system and for different aggregated link data rate.

sum of 90th percentile of throughputs from fronthaul links and the 80th percentile of aggregated link need to be provided (here 200 Mbps). In case of under-dimensioning, for higher percentages of web traffic the delay increase will be lower, as is the sum of the 80th and the 90th percentiles.

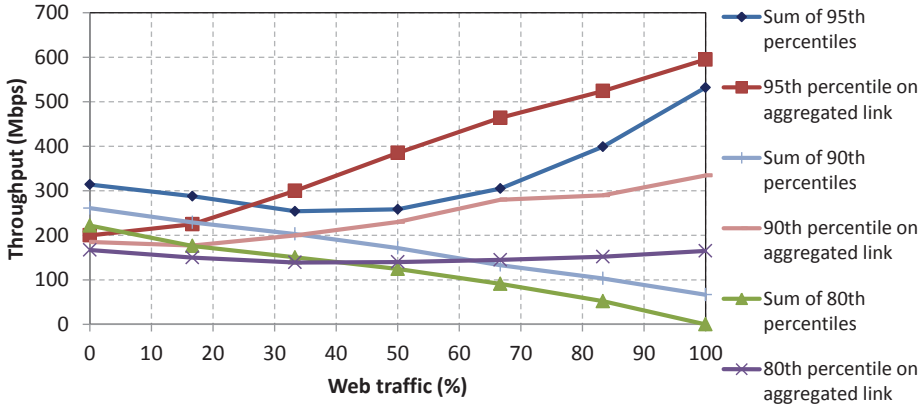


Figure 3.14: 80th, 90th and 95th percentile of base stations throughput for different percentage of web traffic in the system.

These results can be used not only for quantifying multiplexing gains but also for network dimensioning providing traffic distribution with a CDF similar to the one studied here. For upgrading existing networks, operators could measure single cell throughput

and calculate the 90th percentile of it and measure 80th percentile of aggregated traffic. Then depending on how many cells should belong to one BBU pool/be aggregated on a single fronthaul link, the higher number will assure the necessary capacity. If it is not met, it means that links should be upgraded to the next available data rate. For green field deployments based on traffic forecast, the operators will need to estimate what is the 90th percentile of throughput they would like to offer to the users. The challenge then would be to add those throughputs taking into account forecasted user activity. Having the sum of the 90th percentile of such aggregated traffic for each base station, and the 80th percentile of aggregated traffic, the capacities need to be summed considering how many cells will be aggregated on a link/BBU pool. The sum will give the desired link/BBU pool resources.

The results for the sum of the 95th percentiles can be applied to the equation (3.1), where sufficient $AggregatedLinkResources$ are 200 Mbps, based on delay measurements. Using this method computed MG_{95th} is in the range of 1.27 - 2.66 which converges with the results for MG for throughput measured every 10 ms (1.5-2.2). Moreover, 2.7 is the result of $MG_{BBU-BB-RF}$ concluded with the teletraffic method published in [13], which confirms the thesis that $MG_{FH-UE-Cell} = MG_{BBU-BB-RF}$ stated at the beginning of this chapter. These results serve as validation of the two approaches.

3.5 Discussion and verification of the results

A summary of the results achieved in this and another project and impact of averaging, traffic burstiness and tidal effect are shown in Table 3.8. The particular values hold for given application mixes (from the literature) and given daily traffic load in office and residential cells (from an operator - China Mobile as well as a vendor - Ericsson). Still, the trends can be observed and conclusions can be made.

Table 3.8 shows a clear dependence of the averaging period on the multiplexing gain. Depending on the averaging window size, the impact of traffic burstiness and tidal effect can be observed. Sections below discuss those dependencies as well as impact on network model and methodology used to calculate the multiplexing gain. Results are compared thereby allowing to verify network models, assumptions and methodology.

3.5.1 Impact of traffic burstiness

Burstiness is one of the main factors contributing to the multiplexing gain. In order to observe the impact of traffic burstiness, the throughput needs to be measured with fine granularity. Traffic models used in these studies (presented in Tables 3.4 and 3.6) have parameters in order of seconds, e.g., page interarrival time - 10 s, duration of a video - 16 s. Therefore throughput needs to be measured at least every 10 s to observe such traffic fluctuations.

Project 1 shows only impact of traffic burstiness, because 1 hour of network operation was simulated.

Table 3.8: Multiplexing gains calculated in different projects. MG - multiplexing gain.

Project	simulated time (h)	measurement interval	MG		burstiness	tidal effect
1. LTE dimensioning [14]	1	1 s	5.5		✓	-
2. C-RAN MG initial study, Section 3.2.1	24	1 h	1.0 - 1.33		-	✓
3. C-RAN MG initial study [11], Section 3.2.2	24	1 s	4.34 σ 1.42		✓	✓
4. C-RAN MG varying % Office cells [12], Section 3.3	16	576 s	1.2 - 1.6		little	✓
5. C-RAN MG varying % Web Traffic [13], Section 3.4	16	10 ms	1.9 - 2.3		✓	✓
		averaging window W	0% web	≥17% web		
		100 ms	1.6	4.5 - 6.1	✓	✓
		1 s	1.6	3.7 - 5.9	✓	✓
		10 s	1.5 - 3.1		edge	✓
		57.6 s	1.5 - 1.7		-	✓
		100 s	1.5 - 1.6		-	✓

For measurement interval (averaging window) of 1 s in projects 1, 3 and 5 the results are in range of 3 - 6. Therefore this is the value of the multiplexing gain coming from traffic burstiness. Values for project 5, 10 ms averaging are an exception from this reasoning. The values could be lower due to the long tails of CDFs of throughput.

3.5.1.1 Impact of application mix

Project 5 evaluates multiplexing gain for different application mixes varying percentage of web versus video application traffic. Values are the lowest (1.5-1.9) for 0% of web traffic. As soon as the bursty web traffic is present in the network the values increase to 3.7 - 6.1 (project 4, case for 100 ms and 1 s averaging). Similar results are reported for projects 1 and 3 (5.5 and 4.34, respectively). In the majority of networks there is a traffic mix between video and web, e.g., Ericsson reports that in 2015 video accounted for around 50% of mobile data traffic, therefore an operator can count on a multiplexing gain value of 3.7 - 6.1. However, if an operator expects 100% of video traffic (e.g., dedicating the network to users using services like Netflix), multiplexing gain will be lower.

Projects 1, 3 and 5 take into account different application definitions. Still the results remain similar (3 - 6).

3.5.2 Impact of the tidal effect

The tidal effect is another main factor contributing to the multiplexing gain. Section 3.3.2 presents results of three studies that aimed at finding an optimal mix of office to residential cells, all reaching a maximum multiplexing gain for 20-30% of office cells.

Project 4, due to the high value of the averaging window shows mostly the impact of the tidal effect on the value of the multiplexing gain and little impact of traffic burstiness. The value range between 1.2 and 1.6 depending on percentage of office vs residential cells. Figure 3.15 compares values from projects 2 and 4 (latter marked as simulated). Values obtained via simulations in project 4 are 0.2 - 0.3 higher than the ones from initial calculations in project 2 showing little impact of traffic burstiness.

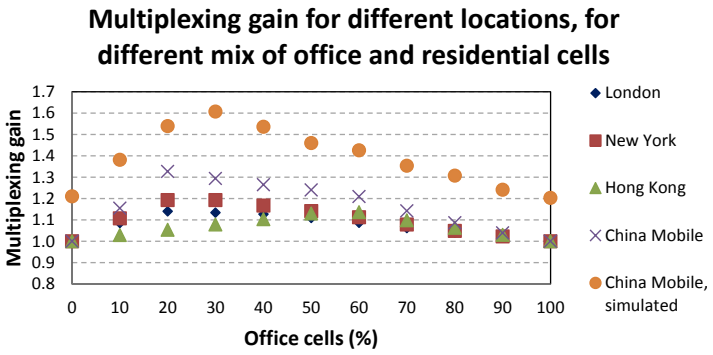


Figure 3.15: Multiplexing gains for different locations based on daily traffic distributions between office and residential cells. Data from China Mobile and Ericsson.

Project 5 explores the most optimal cell mix from project 4. Similar values (1.5 - 1.7) can be observed for project 4, with 57.6 s and 100 s averaging. For longer averaging in project 5, the multiplexing gain value 1.6 matches the value 1.6 from project 4, despite different application definitions and mixes. Therefore it can be concluded that multiplexing gain coming from the tidal effect equals to 1.0 - 1.33. The value is rather small, therefore enabling multiplexing gains should not be treated as the main reason to introduce C-RAN, rather an additional benefit on top of advantages mentioned in Section 2.2. However, there are cells with occasional traffic load, like stadiums, that will greatly benefit from multiplexing gains. Therefore C-RAN is beneficial for such deployments on a local scale, smaller than metropolitan scale.

3.5.3 Impact of the functional split

For BB-RF split multiplexing gain can be observed only on a BBU pool. On the fronthaul links data is constant bit rate, therefore no multiplexing gain can be observed. For the functional splits, where fronthaul traffic is bursty, like UE-Cell split, multiplexing gain can be achieved also on the fronthaul. In both cases values of multiplexing gain reach 2 - 6.

3.5.4 Impact of the model

Value of the multiplexing gain close to 5 is obtained for 1 s averaging for three different projects - 1, 3 and 5 - despite the fact that the first is based on backhaul traffic including whole LTE protocol stack, the second is based on very simple traffic generation and the third covers aggregation of traffic streams on the MAC layer. It is interesting that a very simple traffic aggregation model gave similar results as the advanced LTE model. This similarity shows that for analyzed scenarios protocol processing had little impact on the traffic characteristics. Probably there was little need for e.g. retransmissions.

3.5.5 Impact of the method

All the analysis presented in this chapter (Sections 3.2 - 3.4) is done using a methodology described in Section 3.1 observing properties of aggregated traffic. The tidal effect, application burstiness and protocol processing (for 3.3 - 3.4) are modeled.

In [41], Werthmann *et al.* show results of a simulation study using a different approach - analyzing compute resource utilization, in GOPS, for different number of sectors aggregated into a BBU pool. Dimensioning the hardware according to the 99%-th percentile of the compute load at a system load of 60%, the aggregation of 10 sectors would gain 9%, 20 sectors would gain 20% and 57 sectors would gain 27%, which corresponds to multiplexing gain of 1.09 - 1.37 according to definition presented in equation 3.2. Authors modeled in detail radio network, computation resources as well as locations of users. However, their traffic model does not take the tidal effect into account. It would be interesting to run this traffic definition in model presented in Section 3.4 to compare the baseline scenario, without the tidal effect.

On the contrary, in [40] Namba *et al.* presents an estimation of the multiplexing gain of C-RAN based on the population of the Tokyo metropolitan area. Authors concluded that up to 75% of BBUs can be saved (corresponding to multiplexing gain 4). It is an approximated result as no traffic is taken into account and the assumption is that BBUs correspond to population rather than area the cells cover.

All those three studies are very different in their nature and assumptions. Interestingly, the results presented in Sections 3.2 - 3.4 fit the rage of 1.1 - 4, reaching the value up to 6 only because of traffic profile definition.

3.5.6 Simplifications and assumptions in modeling

As in every modeling work, only some aspects of the system could be captured. C-RAN protocol stack is a complex system, especially when several functional splits are considered. Implementing BBU pool with a full LTE protocol stack is challenging as it requires e.g. complex algorithms for scheduling traffic to many cells and how the processing of this data is handled by various BBUs in the pool, HARQ implementation for the pool. This section aims at listing the assumptions made and how do they limit the results.

3.5.6.1 Protocol stack

The results reported in Section 3.2.2 do not include any protocol processing, while the results reported in Sections 3.2.1, 3.3 and 3.4 include Ethernet protocol processing. All of them resemble variable bit rate data. Actual data for UE-Cell split will be different, because control data will be present and user data will be processed according to the LTE protocol stack. Applied modeling was an approximation in order to be able to use the available tools. OPNET was the key tool used in the thesis. It is an event-based simulation tool that is especially suited to model protocol interactions, however, it does not include physical layer processing on bit level. There is a built-in model of an eNodeB, however PHY layer is not fully implemented there. Moreover, there is also no support for pooling the baseband parts. This model could be potentially interfaced with Matlab, which can model PHY layer processing. However, as the PHY layer performance is implementation specific and involves several parameters, such a model with again be an approximation of a real life system. Additional challenge will arise on how to exactly split layers to model UE-Cell split, how to organize scheduling, HARQ. Moreover, such a complex model will take a lot of time to run. Based on those factors data is generated and sent as Ethernet packets. Such data streams better represent backhaul traffic and it is an approximation to model fronthaul traffic with them.

3.5.6.2 Methodology

In modeling work the traffic aggregation properties were observed. In fact, they refer to the multiplexing gain on transport resources. In order to truly evaluate pooling gain actual computational resources will need to be modeled. Each of the LTE protocol stack layers requires different amount of GOPS, that will scale differently depending on the user activity. Moreover, in C-RAN an additional functionality will need to be in place, e.g. orchestrator, to enable multiplexing gains on BBUs. Full LTE BBU pool implementation will enable a more detailed evaluation of those resources. Lastly, modeling of actual power consumption will give a better insight on possible power savings.

3.5.6.3 Time scale resolution

Traffic measurements in the PhD projects are taken every 10ms - 1h, therefore they do not fully reflect possible traffic peaks that can happen in real time networks. Traffic in LTE is scheduled for each sub-frame - 1 ms. Therefore to observe traffic peaks resources would need to be measured every 1 ms. In project 5., due to the fact that time averaging was applied, the information on traffic peaks is reduced. The results sense what is the maximum value of multiplexing gain if QoS is compromised. The only case where the results could be applicable is if they are measured every few ms - then they would define gains in traffic-dependent resources.

3.5.6.4 Control/user plane

In all the work in PhD project only user plane was considered. Resources needed for control plane processing were not evaluated. Results on $PG_{processing}$ and thereby PG_{power} will be affected when the control plane is included.

3.5.7 Conclusions

Based on the analyzed data set, in a typical, real-life mobile network, with a mix of constant bit rate and bursty traffic, multiplexing gain on traffic-dependent resources in range of 3 - 6 can be expected.

The contribution to the multiplexing gain that C-RAN as such enables on a metropolitan scale is between 1.0 and 1.33, depending on the cell mix. Such a multiplexing gain can be obtained on a BBU pool for all functional splits. To enable higher gains, cells with occasional traffic, like stadiums, should be included in the BBU pool.

However, for the functional splits above UE-Cell, where fronthaul carries variable bit rate data, multiplexing gain 3 - 6 can be achieved on the fronthaul links. This is a strong motivation to investigate and deploy other functional splits than the traditional one, where all the baseband processing is done in the pool.

3.6 Future work

This studies concentrated on multiplexing gain that is dependent on user traffic. As there are parts of base station that need to be on to provide coverage, even when no users are active, it would be beneficial to study overall multiplexing gain including those modules, too.

3.7 Summary

C-RAN is seen as a candidate architecture for 5G mobile networks because of the cost and performance benefits it offers. Due to the fact that securing fronthaul capacity to deploy

fully centralized C-RAN can be costly, a careful analysis of cost and energy savings for different functional splits is of great importance.

Multiplexing gain observed from traffic aggregation approximates possible gains on aggregation transport links. Such a multiplexing gain is proportional to power savings and processing resources savings only on parts of the BBU pool responsible for user-processing functionalities of data-plane processing. However, as exact power and cost models are complex in the analyzed scenarios, the results present only an approximation. Gains in power consumption and pooling gains will be lower than the multiplexing gain.

Pooling gains on processing resources can be achieved only on user-processing resources, which are a fraction of overall signal processing resources - 3-12% on downlink, 17-33% on uplink.

Concerning power savings, 2-24% of total base station power consumption is spend on baseband signal processing. Multiplexing gain can be achieved on user-processing modules, which constitute a fraction of these resources - 3-12% on downlink, 17-33% on uplink.

As only a fraction of resources are impacted by pooling gains, those gains should not be a priority in designing new functional splits. However, a variable bit rate functional split reduces needed bandwidth and enables multiplexing gains on the fronthaul network, which is an important motivation.

Gains in transport are the closest to derive from multiplexing gain value, however, actual traffic patterns in fronthaul network for UC-Cell split will be different than simulated when the LTE protocol stack is included.

Such a lack of the full LTE protocol stack is a major simplification in the modeling work. Others include: methodology based on traffic throughput measurements, not small enough time scale of measurements, and lack of control plane considerations. A full LTE protocol stack BBU pool implementation will enable more precise measurements on pooling gains on baseband resources as well as multiplexing gain on fronthaul link for variable bit rate splits. As energy and cost savings are related to the multiplexing gain, in this study the multiplexing gain is evaluated for different functional splits. Multiplexing gains on throughput-dependent functionalities of a base station are calculated for different C-RAN functional splits: BB-RF and separating user and cell specific functions using four different approaches. For given traffic definitions, a quantitative analysis of multiplexing gains is given. However, the results are sensitive to the traffic profiles as well as to the multiplexing gain calculation method. Therefore the main outcome of this study is to provide the trend lines that will facilitate finding an optimal trade off when fronthaul or BBU resources are more costly for an operator.

For fully centralized C-RAN - with BB-RF split - maximum multiplexing gain on BBU resources can be achieved. However, the required fronthaul capacity is the highest. Therefore this split is vital for operators with access to a cheap fronthaul network. Additionally, if the traffic load is high, the operator will mostly benefit from the multiplexing gain at the BBU pool.

The more functionality is moved from the BBU pool to the cell site, the lower the multiplexing gain on the BBU pool. However, when traffic starts to be variable bit rate, a multiplexing gain on the fronthaul links can be achieved, lowering the required capacity. Hence, for low traffic load, and even more for bursty traffic, the BBU pool should only have higher layer processing and then the requirements on the fronthaul link can be relaxed.

An initial evaluation concludes that up to 1.3 times fewer BBUs are needed for user data processing in C-RAN compared to a traditional RAN looking at daily traffic distributions between office, residential and commercial areas. This number grows to 4 when taking into account specific traffic patterns, making assumptions about the number of base stations serving different types of areas. The latter model does not include mobile standard protocols processing. After including protocols processing the statistical multiplexing gain varied between 1.2 and 1.6 depending on traffic mix, reaching the peak for 30% of office and thereby 70% of residential base stations, enabling savings of 17% - 38%.

The application level delays are verified for different aggregated link bit rates and thereby it is concluded what is the practical value of the multiplexing gain that can be achieved. Rules of thumb for network/BBU dimensioning are proposed based on the CDFs of throughput. The more video traffic is present in the network, the delays are more sensitive to the aggregated link bandwidth, what influences achievable multiplexing gain.

A high impact on the multiplexing gain value is observed depending on the multiplexing gain calculation method, i.e. using different throughput averaging windows. The results vary between 1.0 and 6. The multiplexing gain that C-RAN in traditional, BB-RF split solely enables due to the tidal effect is between 1.0 and 1.33 depending on the cell mix, thereby enabling savings of up to 25%. The major contribution to the numbers higher than 3 comes from traffic burstiness, and in order to benefit from it on the fronthaul links, functional splits that result in variable bit rate on the fronthaul need to be in place.

CHAPTER 4

Towards packet-based fronthaul networks

While the connection between RF and BB parts in a traditional base station was just an interface in one device, in a base station with RRH it often spanned few meters between RRH and BBU in a point-to-point connection, up to the distance from the basement to the rooftop. For C-RAN it is beneficial to connect sites from a metropolitan area, which requires a whole network to support those connections. This network is called fronthaul and spans between cell sites, traditionally equipped with RRH up to the centralized processing location, traditionally equipped with BBU pool.

High capacity requirements on fronthaul are seen as one of the major deal-breakers for introducing C-RAN. Therefore, a thorough analysis of fronthaul requirements as well as solutions to optimally transport data are of high importance, which is reflected in numerous standardization activities, referenced throughout the chapter.

As introduced in Section 2.4, several solutions can be used to organize transport in the fronthaul network, such as: point to point, WDM, WDM over PON, microwave often combined with compression, OTN and Ethernet. This chapter elaborates on the fronthaul requirements (Section 4.1) as well as presents a proof of concept of CPRI/OBSAI over OTN transport (Section 4.2). Moreover, it motivates using Ethernet for fronthaul (Section 4.3) as well as explores challenges and possible solutions to deliver synchronization (Section 4.5 and 4.6) and fulfill delay requirements for a packet-based fronthaul network (Section 4.7 and 4.8). Furthermore, it presents a demonstration of Ethernet fronthaul built in this project (Section 4.9) as well as future directions (Section 4.10). Finally, Section 4.11 summarizes relevant standardization activities, while Section 4.12 summarizes the chapter.

The material presented in Sections 4.3-4.6 was originally published in [15]. It was expended and updated for this dissertation.

4.1 Requirements

This section elaborates on throughput, EVM, delay, jitter and synchronization requirements that fronthaul network needs to support in order for a mobile network to work according to 3GPP specifications.

4.1.1 Throughput

As introduced in Section 2.3.1 a popular LTE cell configuration (2x2 MIMO, 20MHz) will require 2.5 Gbps capacity on the fronthaul link, provided that IQ samples are transmitted using CPRI. For higher bandwidth and more antennas this requirement will scale almost linearly [1]. With the UE-Cell split this throughput can be reduced to 498 Mbps, 700 Mbps [46] or 933 Mbps [148], depending whether 7, 10 or 16 bit sample width is used. [46] assumes 7 bits for DL, 10 bits for UL, while [148] reserves 16 bits. With MAC-PHY bandwidth can be further reduced to 152 Mbps for DL [148].

4.1.2 EVM

Fronthaul network may introduce errors in data transmission. EVM defines how much the received signal is distorted from the ideal constellation points. No matter the fronthaul architecture, general requirements for LTE-A need to be observed for different modulation schemes on the main data bearing channel (Physical Downlink Shared Channel (PDSCH)): Quadrature Phase Shift Keying (QPSK), 16 QAM and 64 QAM, as listed in Table 4.1.

Table 4.1: EVM requirements for LTE-A [[81], Clause 6.5.2]

Modulation	EVM
QPSK	<17.5%
16 QAM	<12.5%
64 QAM	<8.0%

4.1.3 Delay

A HARQ is a process that poses the most stringent delay requirement for LTE-A. As a retransmission mechanism, it takes part in error control and correction. According to the LTE-A standard, for FDD the HARQ RTT Timer is set to 8 subframes (8ms) [170], which means that the user using subframe n needs to know whether retransmission or transmission of new data should occur at subframe $n + 8$, as illustrated in Figure 4.1. Due to the timing advance, the user sends data ahead of time compensating for the propagation delay, in order to fit into the subframe structure at the base station. It appears to be an industry standard that a base station needs to prepare a HARQ acknowledgment (ACK)/non-acknowledgement (NACK) within 3ms [20], [171], [172]: decode UL data, prepare ACK/NACK and create a DL frame with ACK/NACK. Only then will the user receive a ACK/NACK in the 4th subframe after sending the data, 3ms processing time at UE can be accommodated, and a possible retransmission can occur during 8th subframe. If a user will not get ACK/NACK it will retransmit the data in 8th subframe. This 3ms delay budget is spent on BBU and RRH processing as well as UL and DL fronthaul delay,

[1] For CPRI line bit rate option 7A (8110.08 Mbps) and above a more optimal 64B/66B line coding can be used instead to 8B/10B

leaving 100-200 μs [20]/220 μs [46]/ 250 μs [148], [173] for fronthaul one way delay or in other words 200-500 μs Round Trip Time (RTT). Otherwise throughput will be lowered [148]. Section 4.10 discusses possible future directions for the HARQ requirement.

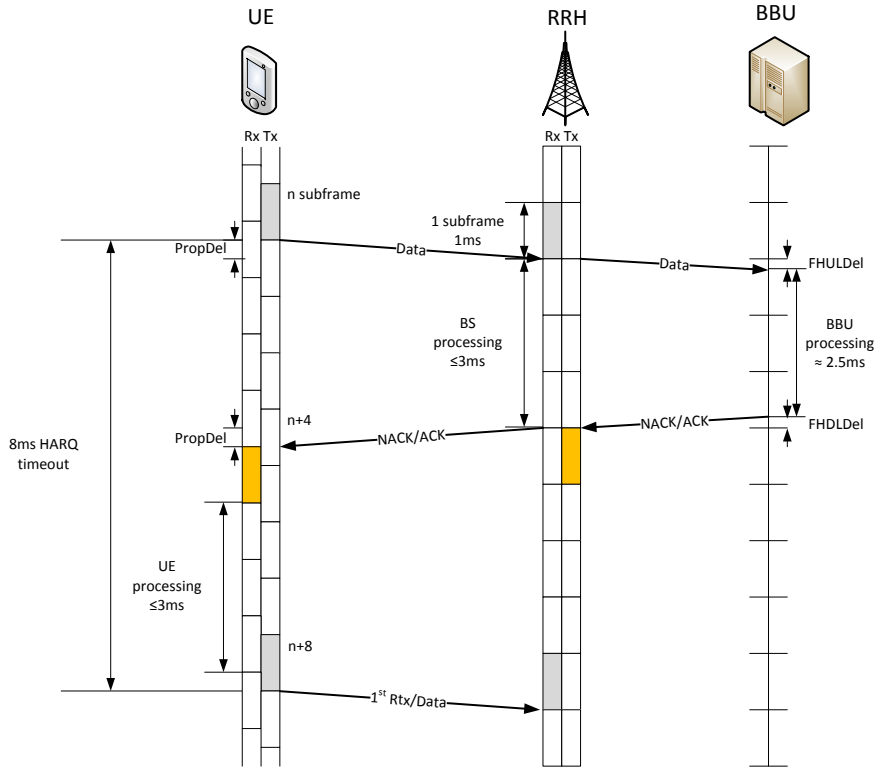


Figure 4.1: Delays associated with the UL HARQ process. PropDel - propagation delay, Rtx - retransmission, FHULDel - Fronthaul UL Delay

Looking at Metro Ethernet Forum Classes of Service for mobile backhaul, Class "High" backhaul could be reused for fronthaul (frame delay $\leqslant 1ms$) [174] or a new class could be specified looking at lower delays, in the order of 100 μs.

4.1.4 Jitter

Nowadays BBUs and RRHs read the delay at the bootup time, therefore the delay needs to be constant. This requirement can be relaxed by buffering the data, however, this will be done at the cost of a higher delay.

4.1.5 Synchronization

A proper synchronization is essential for mobile network operation. In order for a RRH to modulate the data to a particular frequency it needs to know the precise definition of 1 Hz. It is important to keep the carrier frequency sharp in order for the signal coming from the base stations operating in a different frequency band not to overlap and for the UE to be able to receive it. For successful TDD network operation, the RRH needs to follow time frames precisely in order for DL and UL frames not to overlap. Two types of synchronization can be differentiated: frequency and phase (time) synchronization. Clocks are synchronized in frequency if the time between two rising edges of the clock matches. For phase synchronization, rising edges must happen in the same time, as shown in Figure 4.2.

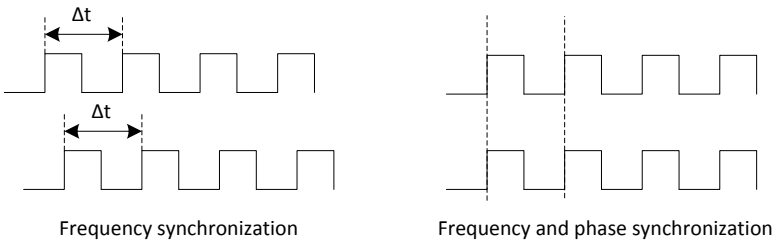


Figure 4.2: Frequency and phase synchronization

In Table 4.2 the requirements that need to be observed for various network features, like Time-Division Duplex (TDD), Frequency-Division Duplex (FDD), MIMO, Carrier Aggregation (CA), eICIC and CoMP are summarized. For latter three, the requirements are expressed relatively to a common reference between the cells/streams involved, otherwise the maximum deviation from an ideal source is listed. Moreover, for Enhanced 9-1-1 services, FCC requires the localization accuracy to be within 50 meters [175], which requires synchronization in the 100 ns range [176].

Taking the example of ± 50 ppb frequency synchronization requirement it is worth noticing that only one-third of this budget, i.e. ± 16 ppb is available for the frequency reference used for the radio frequency synthesis [177], therefore this is the requirement on frequency accuracy delivered over the network. The rest of this budget is consumed within the RRH. The reference requirement of ± 50 ppb dates back to GSM networks [178]. It results from the need to compensate for the Doppler effect for high mobility users and allow for some frequency error tolerance for the receiver at the user side [179]. It is forecasted that for 5G networks this requirement will be even tighter [180].

Table 4.2: Summary of timing requirements for LTE. BS - base station

Feature	Frequency	Time
LTE-A FDD	±50 ppb (wide area BS) ±100 ppb (local area BS) ±250 ppb (home BS) [[81], Clause 6.5.1.1]	–
LTE-A TDD		±5 μs (cell with radius $> 3km$) ±1.5 μs (cell with radius $\leq 3km$) [[82], Clause 7.4.2]
MIMO, CA		$\leq 65ns$ [[81], Clause 6.5.3.1]
eICIC		±1 μs [181]
CoMP		±1.5 μs [181]
CPRI	±2 ppb [Req. R-18 [33]]	±16.276 ns [Req. R-20 [33]]
911 call localization		100 ns [176]

4.2 OTN-based fronthaul

OTN can be found in many fiber network deployments and can be reused to enable C-RANs. Mapping of CPRI/OBSAI bit streams over OTN is done over low level ODU containers as specified in ITU-T G.709/Y.1331. The main challenges of transmitting CPRI/OBSAI over OTN is to limit the frequency error introduced while mapping and de-mapping CPRI/OBSAI to OTN. Moreover, the signal must not be deteriorated to maintain its characteristics. Deterministic awareness of the delays through each hop needs to be maintained (not covered in this work).

- OTN is a standard proposed to provide a way of supervising client's signals, assure reliability as well as achieve carrier grade of service.
- OTN is a promising solution for optical transport network of C-RAN when existing OTN legacy network can be reused for C-RAN fronthaul connecting RRHs to the BBU Pool.
- OTN provides FEC allowing the transport of client signals like CPRI in noisy environments or over longer distances.
- OTN supports both wavelength- and time-domain multiplexing maximizing the bandwidth utilization of the fiber network.

OTN-based fronthaul was experimentally verified by connecting a Base Station Emulator (BSE) and a RRH provided from MTI Radiocomp to an OTN compliant client-to-OTU2 Mapper and Multiplexer from Altera using both CPRI and OBSAI bit streams at different rates (2.4576 Gbps and 3.072 Gbps, respectively) carrying LTE traffic. It was observed that carrier Frequency Error and EVM changes are within Third Generation

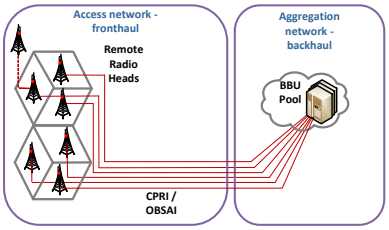
Partnership Project (3GPP) specifications for LTE-Advanced. It was a very useful exercise towards understanding frequency error and EVM. More details about those measurements can be found in Appendix A.

4.3 Motivation for using Ethernet-based fronthaul

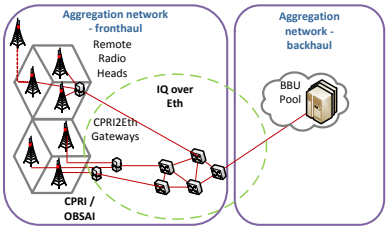
The traditional view on the C-RAN architecture is to connect multiple RRHs to the BBU Pool, each using a dedicated fiber connection to carry IQ data, as shown in Figure 4.3(a). Traffic is delivered using a constant bit rate protocol - CPRI, or the less popular OBSAI. Following up on author's earlier work presented in Chapter 3, in order to decrease the cost of fronthaul network and a BBU pool, a packet-based fronthaul was proposed, where variable bit rate data from the new functional split can be transported in packets, as well as cells from residential and office areas can be dynamically associated to different pools to maximize multiplexing gain. The proposed architecture is shown in Figure 4.3(b). In order to further optimize the cost, fronthaul network could utilize widely-deployed existing Ethernet deployments so that IQ data shares resources with other types of traffic. Moreover, Ethernet enables traffic aggregation and switching and should be cheaper than e.g. deploying WDM networks.

Packet-based fronthaul architecture is of high interest for industry.

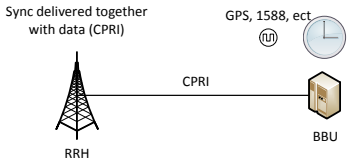
- NGMN performed initial analysis of functional splits, which included variable bit rate fronthaul, which can be most optimally transported in packet-based manner [68].
- NGFI working group is under preparation at IEEE [180] aiming at standardizing new functional split decoupling the fronthaul bandwidth from the number of antennas, decoupling cell/UE processing, focusing on high-performance-gain collaborative technologies as well as reliable synchronization and data transport in packetized networks [182], [46].
- In October 2014 the 1904.3 Task Force [183] has been started in IEEE to standardize a transport protocol and an encapsulation format for transporting time-sensitive wireless (cellular) radio related application streams in Ethernet-based transport networks. The new standard will be called "Standard for Radio over Ethernet Encapsulations and Mappings". Both CPRI and any other functional split can be supported. The standard is also independent on the synchronization solution. The scope of the standard is as follows:
  - encapsulation of IQ data to Ethernet packets, both in terms of a header format and actual CPRI streams to Ethernet packets mapping,
  - defining a protocol that enables multiplexing of several streams,
  - defining a protocol that supports sequence numbers/timestamps to allow the received data to be played back keeping original timing,



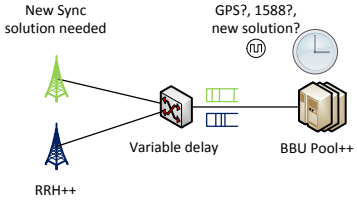
(a) Traditional C-RAN architecture with a dedicated line to each RRH.



(b) C-RAN with packet based fronthaul.



(c) Synchronization in a traditional fronthaul.



(d) Synchronization in a packet-based fronthaul.

Figure 4.3: Traditional and discussed C-RAN architecture together with principles of deriving synchronization for them

- a control protocol for link and endpoint management.

The use cases of IEEE 1904.3 cover both legacy RRH and BBU as well as Ethernet-based RRH and BBU. The first version of the standard is planned for May 2017. IEEE 1904.3 will be referred to as 1904.3 later on in the text.

- CCSA founded a project to study the requirements, scenarios and the key technologies for next-generation fronthaul [180].

- Fronthaul is an essential topic in recently founded ITU-T IMT-2020 Focus Group, which looks into how to support synchronization for NGFI [180].

4.4 Challenges of using packet-based fronthaul

In the traditional C-RAN architecture synchronization and timely delivery of traffic are ensured by using a synchronous protocol - CPRI, or the less popular OBSAI. Today's mobile technologies, especially LTE-A, require high accuracy in terms of frequency and phase for proper transmission. By using C-RAN, these requirements are extended to the link connecting the RRH and the BBU Pool, and thereby to the Ethernet. It is not in the nature of Ethernet to be synchronous, which conflicts with CPRI. The main challenge is then to find a method of assuring synchronization and latency across Ethernet meeting the demands of LTE-A, which this section focuses on.

Secondly, in packet-based network queues in the switches can be expected which makes it challenging to support tight delay and jitter requirements detailed in Section 4.1.

Sections below explain those challenges in more detail as well as propose solutions to address them.

4.5 Technical solutions for time and frequency delivery

4.5.1 Synchronization today

In current networks BBU is typically equipped with a high precision clock having the time source in form of Global Positioning System (GPS) or IEEE 1588 Precision Time Protocol (PTP) [184] possibly supported by SyncE, as shown in Figure 4.3(c). Currently deployed, and considered in this dissertation, is the second version (v2) of the protocol, also referred to as IEEE-2008. Later in the text, IEEE 1588 v2 will be referred to as 1588. GPS and 1588 can deliver both frequency and phase synchronization, while SyncE can only deliver frequency synchronization. Therefore, SyncE is often used as a complementary solution to GPS or IEEE 1588 to enhance frequency accuracy. RRH gets precise time information via CPRI link, where timing information is included together with a data stream.

4.5.2 Synchronization in the future

In the future, a fronthaul link might be multipoint-to-multipoint with packet-based transport which will affect synchronization. A comparison between the synchronization solution used today and a scenario for the future is illustrated in Figure 4.3. Especially challenging in Ethernet networks is the fact that packets will experience variable delays passing through the switches. Moreover, the Ethernet switches themselves have uncertainty of their internal clocks in the order of $\pm 100 ppm$ [[185], clause 22.2.2.1] and introduce a timestamping error, that influences 1588 performance. The timestamping error depends on the switch clock frequency and the delay introduced in the de-serialization circuit, and can be reduced, if the delay in the de-serialization circuit can be estimated.

To use legacy RRHs, a CPRI2Eth gateway is needed to bridge the Ethernet and CPRI domains. For future RRHs (RRHs++) the Ethernet link could terminate at the RRH, omitting CPRI. A possible solution for delivering synchronization is to equip the

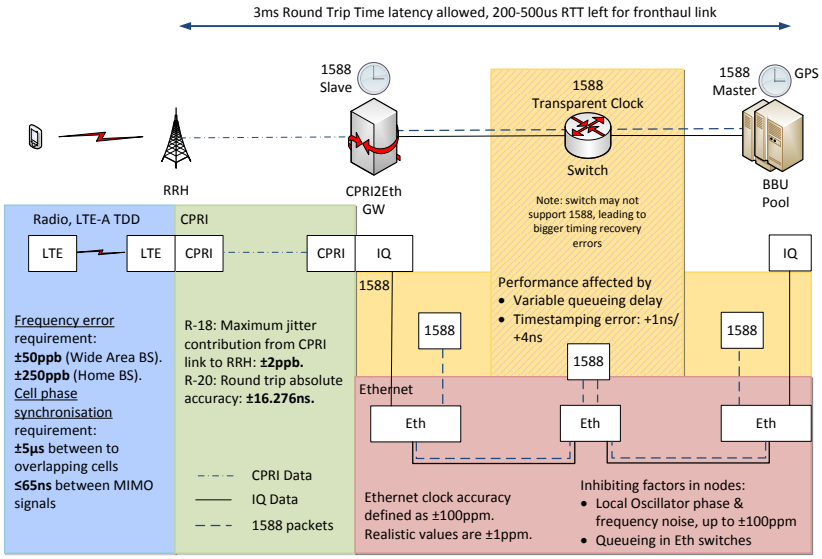


Figure 4.4: Model of the requirements and factor introducing uncertainties in LTE, CPRI, 1588 and Ethernet layers.

CPRI2Eth gateways or the RRHs++ with GPS. This solution assures both frequency and phase delivery. However, it increases the cost, spent not only on GPS equipment but also on an oven-controlled oscillator. Moreover, the coverage indoors and in metropolitan valleys (small cell on a lamp post between high buildings) will be problematic. For some operators it is also important not to depend on a third-party solution for their network. Another solution could be to implement a 1588 slave in the CPRI2Eth gateways or in the RRH++. This solution assures lower equipment cost, however, it will be affected by variable network delay present in Ethernet networks. Ashwood [186] shows that such a jitter, when the background traffic is present, can be in the order of μs per Ethernet switch.

The architecture considered is presented in Figure 4.4. Ethernet packets are sent from the BBU through the network of switches to reach the CPRI2Eth gateway. They are repacked there to CPRI stream to reach legacy RRHs. Alternatively, Ethernet RRHs can be used, omitting CPRI. A packet-based solution - IEEE 1588 - is used to assure synchronization with a 1588 master present in BBU Pool and a 1588 slave present in CPRI2Eth gateway. Figure 4.4 summarizes requirements on different layers: LTE, CPRI and Ethernet as well as factors introducing inaccuracies in 1588 and Ethernet layers. In the section below a feasibility study of using 1588 for timing delivery is performed. The modeling work takes into account the factors influencing the performance of 1588

mentioned in Figure 4.4. A dedicated Ethernet network was taken into account, leaving for future work the case of sharing Ethernet infrastructure with other types of traffic.

It is worth noticing, that the requirements need to be fulfilled on two levels:

- delay requirements,
- frequency accuracy requirements,

where the delay of synchronization packets lies between them, as 1588 helps to recover the clock, but is affected by network delays, as emphasized in Figure 4.5.

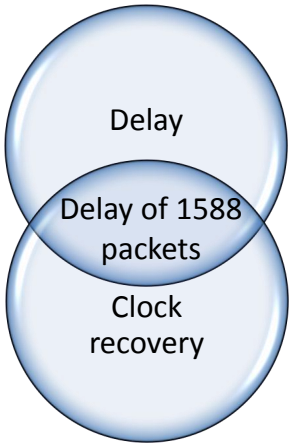


Figure 4.5: Time related requirements on a fronthaul network

4.6 Feasibility study IEEE 1588v2 for assuring synchronization

The 1588 standard defines a set of messages ($Sync$, $DelayReq$ (delay request) and $DelayResp$ (delay response)) for an end-to-end operation in order to exchange a set of timestamps between master and slave clocks as shown in Figure 4.6. The master clock needs a precise time definition e.g. from GPS and then propagates this timing information to the slaves. Timestamp t_1 is inserted into the $Sync$ message when it leaves the master node, t_2 is noted when the message arrives to the slave. In the model, each time $Sync$ and $DelayReq$ messages pass through an Ethernet switch (working as a 1588 Transparent Clock (TC)) the value of $CorrectionField$ (CF_S and CF_D, respectively) is updated for the residence time as presented in the Equation (4.1). Recording residence time is very important as variable traffic in packet-based networks will fill-in the queues leading to variable network delay for 1588 packets, thereby affecting time reference arriving at the slave. In response to the $Sync$ message slave sends the $DelayReq$ noting down when the

message leaves the node - timestamp t_3. The master node notes the time when it receives the message - t_4 and sends it to the slave via the $DelayResp$ message. Based on those timestamps, the delay and offset between the clocks can be computed as shown in the Equation (4.2) and (4.3), respectively. It can be observed that the messaging of the time stamps from the 1588 master to the slave is a feed forward messaging algorithm. More information on the 1588 operation can be found in [187].

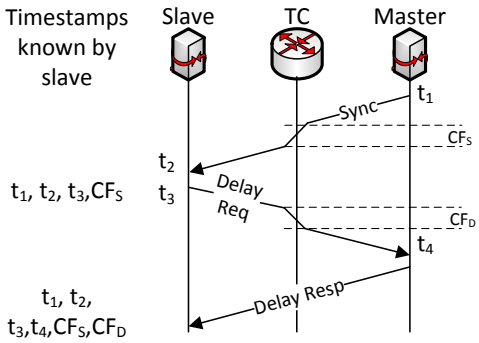


Figure 4.6: Visual representation of 1588 operation.

$$CF = EgressTimestamp - IngressTimestamp \tag{4.1}$$

$$Delay = \frac{(t_2 - t_1) + (t_4 - t_3) - CF_S - CF_D}{2} \tag{4.2}$$

$$Offset = \frac{(t_2 - t_1) - (t_4 - t_3) - CF_S + CF_D}{2} \tag{4.3}$$

A protocol stack of the suggested solution is shown in Figure 4.7. Here the relevant protocols when CPRI traffic is transmitted over Ethernet using 1588 are shown.

The crucial aspect in implementing 1588 functionalities is to execute timestamps generation as close as possible to the moment when each packet enters/leaves each node. It is important that the timestamp t_1 is taken exactly when the $Sync$ message leaves the Master node, t_2 when the $Sync$ message enters the Slave node etc. Otherwise, the variable internal processing time of the packets will affect the measurements. It is especially important in case of Ethernet switches, as a variable residence time is expected, depending on variable traffic queuing up in the switches. Therefore, for the correction fields ingress and egress timestamps should be taken as soon as the packet enters and leaves the node, respectively, as shown in Figure 4.8 in order to compute the $CorrectionField$ value precisely.

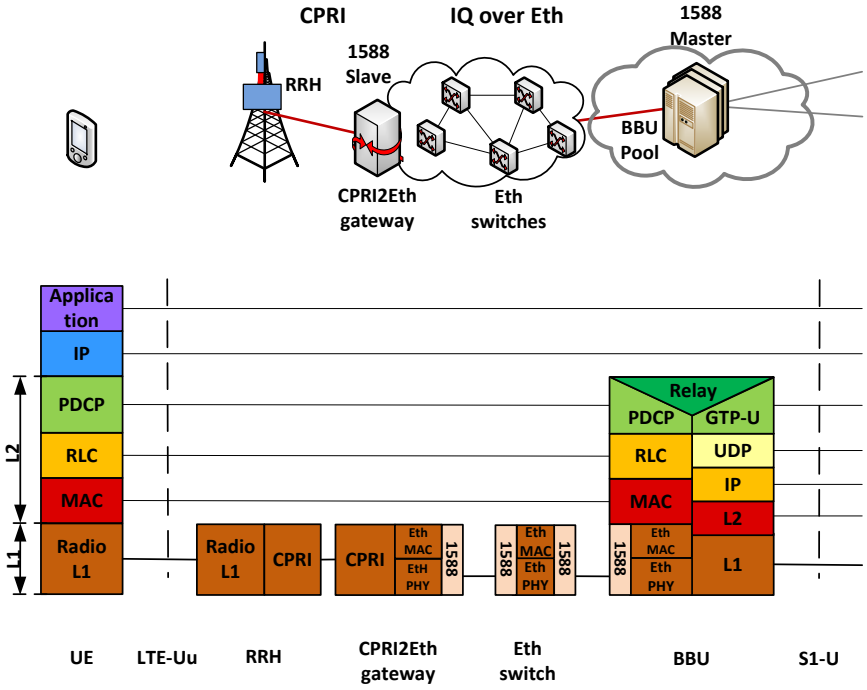


Figure 4.7: Protocol stack of the examined network.

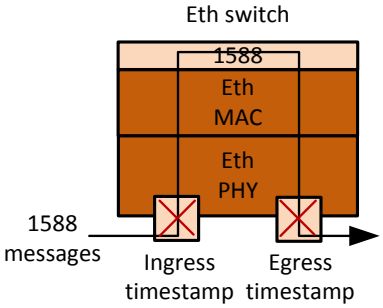


Figure 4.8: Ingress and egress timestamps should be taken as soon as $Sync$ or $DelayReq$ packets enter and leave the node, respectively.

The 1588 standard defines the establishment of a clock hierarchy and format of messages, so that the equipment from different vendors can communicate in the network. The standard gives an example on how to calculate the clock drift as presented in the Equation (4.4). However, the implementation of an actual synchronization algorithm lies outside of its scope and this can possibly give a competitive advantage to certain vendors over others in performance of their solutions. Various works have been published presenting synchronization algorithms. Xie *et al.* in [188] propose to maintain two time scales in the slave: syntonous (frequency aligned) and synchronous (time and frequency aligned). For delay calculations they use t_2 and t_3 measured in a syntonous time scale. They concluded that it is the most optimal to apply drift correction every third time the exchange of 1588 messages is completed. This way of drift calculation was implemented, as in Equation (4.5).

$$DriftStd = \\ = \frac{t_{2(N)} - t_{2(0)}}{(t_1 + Delay + CF_S)_N - (t_1 + Delay + CF_S)_{(0)}} \tag{4.4}$$

$$DriftImplemented = \\ = \frac{t_{2(N)} - t_{2(N-3)}}{(t_1 + Delay + CF_S)_N - (t_1 + Delay + CF_S)_{(N-3)}} \tag{4.5}$$

A network model in OPNET modeler was built checking the performance of this algorithm. OPNET is an event-driven simulation software, where a user can build his scenario from self-defined nodes and processes. A network consisting of a 1588 Master, a 1588 Slave and a variable number of Ethernet switches working as 1588 transparent clocks was built. $Sync$ and $DelayReq$ packet rate is 64 packets per second (pps). The novelty of this work is a network view where the protocol was tested against various errors that can occur in the network. The slave node has an initial frequency drift of 1 ppm or 100 ppm (maximum that an Ethernet switch can have). Each of the Ethernet switches has a frequency error of 1 ppm or 100 ppm and timestamping error of 1 ns or 4 ns. As it is not possible to measure the exact arrival time of a packet using the internal clock reference in the switch, a random timestamping error is introduced up to 1 ns and 4 ns, respectively. The values mentioned above represent parameters of newer and older generation equipment from the industry. The models follow the protocol stack from Figure 4.7 on the link between the BBU and the CPRI2Eth gateways. 2.5 Gbps CPRI traffic was sent over a 10 Gbps Ethernet network. In between master and slave node 0-21 Ethernet switches were put. The simulations were run for 10 minutes, while CPRI traffic was sent after 30 s. 30 seconds were required to get the network operational with Ethernet switching topology getting established.

For each exchange of 1588 messages, after all timestamps are gathered, delay, drift and offset are computed. Drift correction is applied to slave clock frequency f_S every

third exchange of timestamps. That affects both synchronous and syntonous time scales at slave having impact on $t_{2(Syntonous)}$ and $t_{3(Syntonous)}$. The synchronous time scale of the slave is updated for the offset after each exchange of timestamps. That affects local time at the slave t_S. A relative frequency error between master clock frequency f_M and slave clock frequency f_S was measured, as presented in Equation (4.6) and an absolute phase error between time in master t_M and time at slave t_S as presented in Equation (4.7).

$$FrequencyError = \frac{f_M - f_S}{f_M} \tag{4.6}$$

$$PhaseError = t_M - t_S \tag{4.7}$$

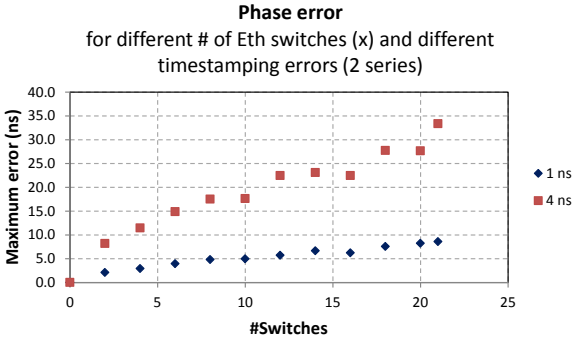


Figure 4.9: Maximum phase error measured for various scenarios during stable operation.

Figure 4.9 presents the maximum observed phase (time) error during stable operation (after initial time discovery) for different numbers of the Ethernet switches present in the network and two different timestamping error values. The phase error stays in the order of nanoseconds and is highly dependent on timestamping errors. The dependency is close to linear. The results are shown for the worst-case scenario of 100 ppm drift for both the Ethernet switches and slave, as drift value had marginal effect whether it was 1 ppm or 100 ppm in both cases.

Figure 4.10 shows the frequency error for the afore mentioned scenarios. The frequency error falls way above required values (16 ppm or below, depending on implementation). It is also highly dependent on timestamping errors. That is the reason why the improvements to this method were applied.

4.6.1 Improving the 1588 slave design

In order to improve 1588 performance, averaging to both offset and drift correction was applied. In that way if one of the packets experienced delay way longer than others, due to e.g. queuing, this outrageous measurement will have lower impact on the slave clock adjustments. The method is presented in Equation (4.8) for the drift and applies also for

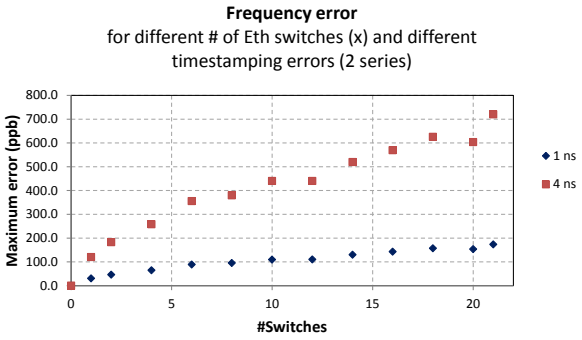


Figure 4.10: Maximum frequency error measured for various scenarios during stable operation.

the offset. The frequency is adjusted by averaged drift computed taking only a fractional value of currently computed drift - $drift$ - and previously computed drift - $driftPrev$. The performance of the system for different values of $alpha$ was checked, and it was concluded that the higher it gets, the lower frequency error is observed. However, for higher values of $alpha$ it takes more time for the system to converge to stable operation. For simulations $alpha = 0.99$ was used. After 180 s the system reached stable operation. The frequency error got significantly smaller for all the cases (20 times smaller), while the phase error got slightly smaller (2 times) as presented in Figures 4.11 and 4.12.

$$driftAvg = alpha \cdot driftPrev + (1 - alpha) \cdot drift \tag{4.8}$$

$$freq = freq - driftAvg \tag{4.9}$$

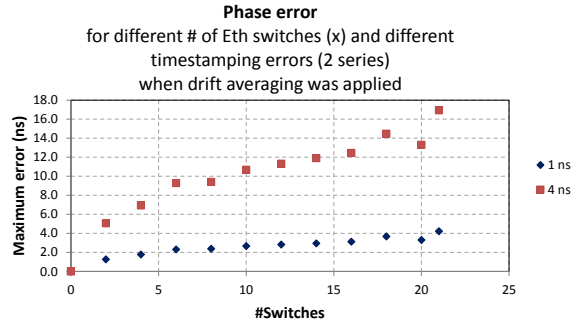


Figure 4.11: Maximum phase error observed during stable operation for various scenarios with offset averaging applied.

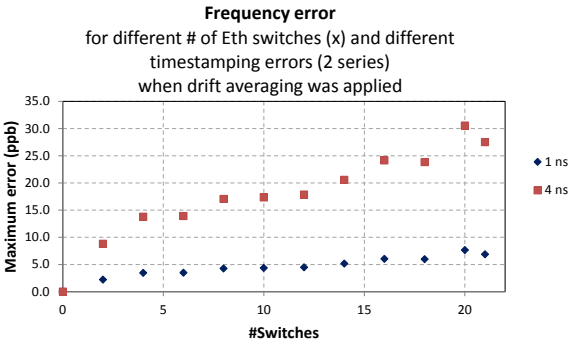


Figure 4.12: Maximum frequency error observed during stable operation for various scenarios with drift averaging applied.

4.6.2 Presented results vs mobile network requirements

It should be noted what the dependency between synchronization inaccuracies of a local oscillator and the inaccuracies of an RF signal is. As presented in Figure 4.13, an Ethernet signal entering an RRH carries both the IQ baseband signal and 1588 packets carrying timestamps. The 1588 module processes timestamps from $Sync$, $DelReq$ and $DelResp$ messages in order to calculate drift and offset of a clock. This information serves as an input to the Phase-locked Loop (PLL) system. The PLL then controls the local oscillator. The baseband signal is processed to become an RF signal and the timing information is taken from the local oscillator. In Section 4.1, requirements for an LTE-A RF signal are outlined. In this work, the accuracy of 1588 inputs are studied. It is implementation-specific how this will affect the PLL performance, the local oscillator and as a consequence, the inaccuracies of an RF signal. These factors should be taken into account designing the clock recovery mechanisms. For the frequency up conversion from the local oscillator f_{LO} to the RF frequency f_{RF}, a scaling factor $\frac{f_{RF}}{f_{LO}}$ will proportionally scale the frequency error (in the example in Figure 4.13 it is equal to 10).

4.7 Technical solutions for delay and jitter minimization

The sections above focus on evaluating synchronization challenges and possibility of applying IEEE 1588 to deliver clocking reference. On top of synchronization requirements, as stated in Section 4.1, fronthaul network needs to be able to transmit the data within 100-250 μs, depending on BBU processing time. In this Section various factors affecting the delay are analyzed. Moreover, two mechanisms: source scheduling and preemption are presented, that aim at reducing delay and jitter. In the next Section a design of source scheduling algorithm and considerations on its implementation are presented.

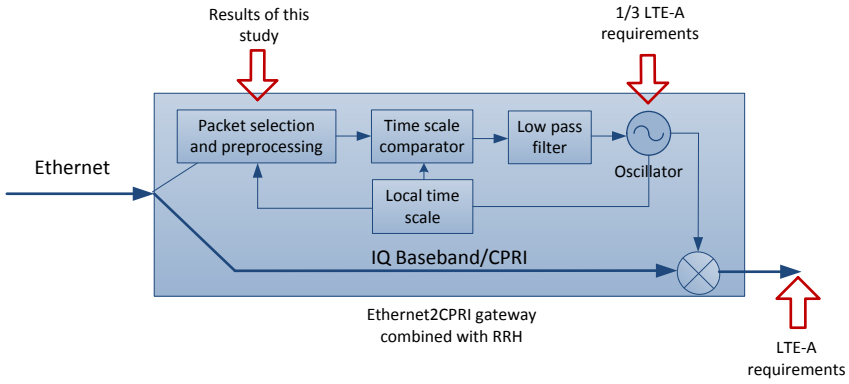


Figure 4.13: Clock recovery scheme inside an RRH combined with CPRI2Eth gateway. LO - local oscillator

Table 4.3: Delays in an Ethernet switch

	Maximum Transmission Delay 10 Gbps link, μs				Switching Delay, μs
Number of switches	MTU, B				E.g.
	1500	3000	6000	9000	
1	1.2	2.4	4.8	7.2	3.0
2	2.4	4.8	9.6	14.4	6.0
5	6.0	12.0	24.0	36.0	15.0
10	12.0	24.0	48.0	72.0	30.0

4.7.1 Analysis of the delay in Ethernet network

The main challenge in enabling an Ethernet-based fronthaul is to keep the queuing delay as low and with as little jitter as possible.

The network delay has the following components, as illustrated in Figure 4.14:

- propagation delay – deterministic, 5 $\mu s/km$,
- switching delay – deterministic, for store-and-forward switches it is in the μs order of magnitude, e.g. 3 μs in the switches measured, as presented in Section 4.9; for cut through switches will be shorter, in order of the ns,
- transmission delay – deterministic, depending on the packet size and link speed, examples for 10 Gbps Ethernet link are shown in Table 4.3,
- queuing delay – non-deterministic, depending on the load of the network.

The examples of delay budgets are defined in Table 4.4. Figure 4.15 shows the dependency between RRH-BBU distance, number and type of switches as well as MTU

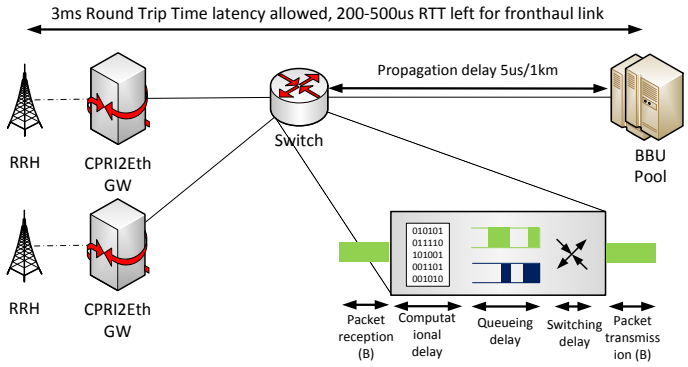


Figure 4.14: Delays in Ethernet-based fronthaul

Table 4.4: Exemplary delay budgets

#	Total distance, km	# switches	Type of switch	MTU	Delay on 10 Gbps link, μs
1	20	10	store-and-forward	9000	100 + 30 + 72 + queuing = 202 + queuing
2	10	10	cut-through	9000	50 + ns + 72 + queuing = 122 + queuing
3	10	5	cut-through	1500	50 + ns + 6 + queuing = 56 + queuing

size. It assumes no queuing delay. Figure 4.16 takes the best case (cut through switches, MTU 1500) and evaluates it for different queuing delay per switch.

If a dedicated Ethernet link is used for a fronthaul link, there will not be any queuing delay. However, for multiplexing gains on links it is desired that many fronthaul streams will share the link, possibly also with other types of traffic. Even if various fronthaul streams will be given the highest priority, it needs to be assured they will not collide. Moreover, the lower priority packets should not slow down fronthaul packets if they happen to be under processing by a switch. The following two methods are aiming to address these problems and are currently under the standardization in IEEE:

1. Traffic scheduling - IEEE 802.1Qbv Bridges and Bridged Networks — Amendment: Enhancements for Scheduled Traffic [189]

2. Preemption – IEEE 802.1Qbu “Bridges and Bridged Networks – Amendment: Frame Preemption” [190]

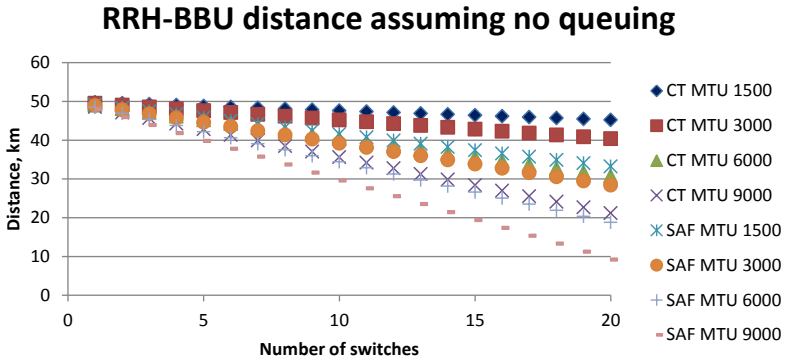


Figure 4.15: Allowed distance between RRH and BBU for a total delay budget of 250 μs depending on the number of switches in the network, MTU size and type of switch (SAF - Store and forward, CT - Cut through). Assumed queuing delay is zero.

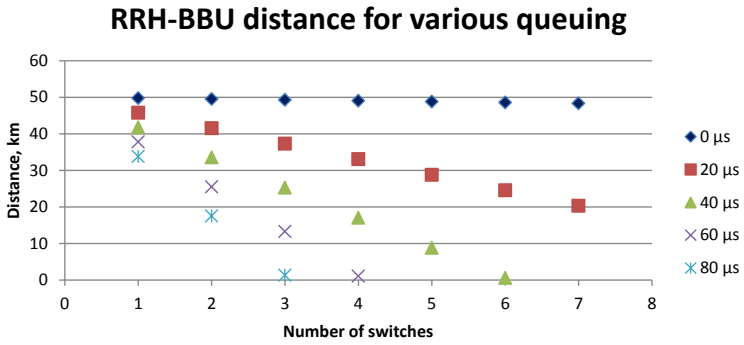


Figure 4.16: Allowed distance between RRH and BBU for a total delay budget of 250 μs depending on the number of switches in the network, MTU size and queuing delay per switch. CT switch, MTU 1500.

Both solutions belong to the TSN set of standards, which can be applied to any time sensitive application. In April 2015 the Project Authorization Request (PAR) of a new standard was prepared - 802.1CM Time-Sensitive Networking for Fronthaul that aims at defining profiles that select features, options, configurations, defaults, protocols and procedures of bridges, stations and LANs for the fronthaul applications [191]. The sections below describe the source scheduling and preemption.

4.7.2 Lowering jitter using source scheduling

When the fronthaul data needs to share the link with traffic with different priority, it needs to be assured that they will not collide. Moreover, when the Ethernet links will be shared between many fronthaul streams it needs to be assured that they will not collide, either. In order to keep the fronthaul delay deterministic, it might be beneficial to delay the time when a single packet leaves the source, so that it will not collide with other packets in the network. The ideas behind traffic scheduling are standardized in the IEEE 802.1Qbv specification. However, they can be implemented by using any type of controllers e.g. SDN or control packets of IEEE 1903.4.

Case 1: one fronthaul stream (protected traffic), one or more other traffic streams (unprotected traffic) One approach to achieve that can be to assure that only one stream has access to the network at specific times (protected window, from T1 to T2 in Figure 4.17), in other words, the transmission for protected traffic is scheduled between time T1 and T2. However, in order to make sure that the unprotected traffic is not under transmission while the protected one arrives, a certain guard band needs to be in place (T0-T1). The simplest solution would be to have a guard band equal to the transmission time of the largest packet size supported by the network. That however, leads to resource inefficiency, as the network could be unnecessarily idle. More optimally, the implementation could check if there are any packets in the queues, whose transmission could end before the transmission of protected traffic. Unprotected traffic will be sent on the best effort basis.

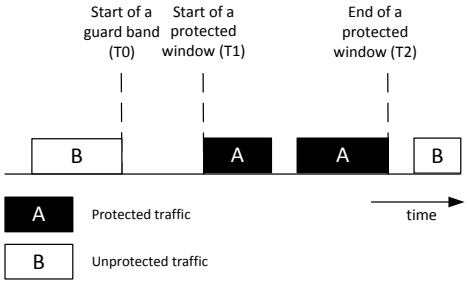


Figure 4.17: Protected window, here for the fronthaul traffic

Case 2: many fronthaul streams In this case it is not sufficient to schedule a protected window when the fronthaul stream requests it, as many streams may collide, creating variable (and non-deterministic) delay due to queuing in the switches. In the fronthaul application it is important that the delay would be as stable as possible (low jitter) while maintaining the delay within the requirements. Traffic scheduling can be implemented already in the sources and it is illustrated in Figure 4.18. Packets A and B initially would arrive at the switch at the same time and one of them will need to wait in the queue, therefore experiencing a non-deterministic delay. When one of the packets – here B - is initially delayed its delay would be larger, but deterministic-predictable. It is especially important when packets are going through many switches and a non-deterministic delay would create a big jitter.

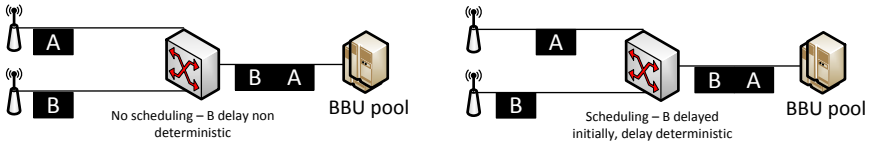


Figure 4.18: Source scheduling used to reduce jitter. Here an example for UL

Case 3: many fronthaul streams and one or more other traffic streams (unprotected traffic) Traffic scheduling as described in the section above can be used between many fronthaul streams. Unprotected traffic will be sent on the best effort basis.

Using gate opening to create protected windows The 802.1Qbv standard allows the implementation of traffic scheduling by means of a gate opening. A sequence of gate open/close operations can be scheduled for each port and traffic class, allowing each traffic class to be sent at a given time but be blocked at others. The schedule can be changed periodically. A frame of a given class can be transmitted only if the gate is open and it will remain open long enough for the packet to be transmitted.

4.7.3 Lowering delay and jitter using preemption

Even if the packet has the highest priority, when it arrives at the switch there might be another lower priority packet already under processing. In case jumbo frames are used, the delay might be up to 7.2 μs on a 10 Gbps link. It is shown in the case for no preemption in Figure 4.19. The low priority packet B is transmitted first and only then can the high priority packet A be transmitted.

This problem can be addressed with preemption. With preemption a lower priority – preemptable – packet can be interrupted from a transmission when a high priority packet arrives. This scenario is show in a Figure 4.19: packet B is preempted, packet A transmitted and then transmission on packet B can be resumed. At the receiver end,

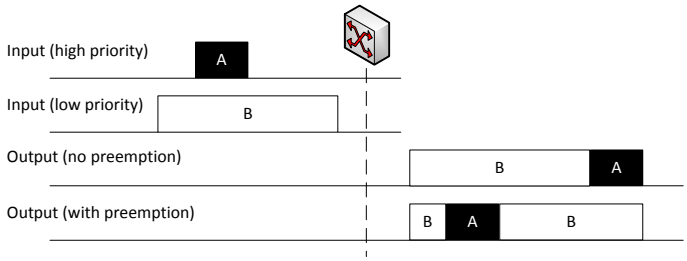


Figure 4.19: Preemption

preempted packets are then merged. The technical details of preemption are described in 802.3br Draft Standard for Ethernet Amendment: Specification and Management Parameters for Interspersing Express Traffic [192].

This approach is especially useful when a fronthaul stream will share the network with other applications. It can optimize guard band size for scheduled traffic, as bigger parts of the packets can fit before gate closing. The smallest size of a preempted packet needs to be not smaller than 64 B, and thereby packets up to 124 B can be preempted. Alternatively, only with source scheduling the whole packets will need to wait for transmission, even if the switch would be free for some time. Therefore the capacity would be wasted. With preemption it is safe to allow a portion of the packet to be sent.

4.7.4 Discussion

As can be seen from the delay analysis, it is possible to enable Ethernet-based fronthaul without TSN functionalities for dedicated links. However, as soon as more than one fronthaul stream will use the link, it is important to observe potential queuing delays. Source scheduling is a beneficial technique that requires intelligence only at the network edges (CPRI2Eth gateway or Ethernet RRH), while the legacy switches can be used. If other services are supposed to share the network with the fronthaul streams, then the delay may get even bigger. Preemption is a technique that assures timely delivery of higher priority packets at the same time allowing other traffic to use the link whenever it is not colliding with fronthaul packets.

If only preemption is enabled, than the delay can be kept deterministic for only one stream. For more streams algorithms minimizing the overall delay on the network level would need to be in place.

For fast networks (10+ Gbps) with low load TSN features are less significant.

Wan *et al.* in [193] presents discrete-event based simulation results on transmitting CPRI over Ethernet. The authors measured that for a tree topology network where each traffic stream went through 2-4 switches the delay was about 90 μs and the jitter was up to 400 μs. When background traffic was inserted, the Ethernet with preemption performed similarly - delay was 91 μs and the jitter was 410 μs. However, with the scheduling algorithm the authors proposed the performance was not consistent - in the majority of the

cases the jitter was removed, but in some grew up to 1000 μs. The authors recommend to use Ethernet with preemption and buffering at the edge.

Farkas *et al.* in [194] present other simulation results of CPRI over Ethernet transport with TSN features. Without the TSN features in a tree topology comprised of 10 Gbps links, where each stream went through 1-3 switches, the switching delay was 1500 ns $\pm$ 5 ns variation. For the scenario with background traffic, the jitter raised up to 3 μs, but when the preemption was enabled jitter stabilized in 100 ns level. With source scheduling the jitter amounted to up to 50 ns. When source scheduling with 70 ns guard band was added on top of preemption the jitter was reduced to 0 ns. The delay for all the cases was below 26 μs. Therefore it was shown in simulations that usage TSN standards and a proper network configuration the delay and jitter requirements for fronthaul can be met.

4.8 Source scheduling design

The previous section describes two approaches that are under standardization and can be used to reduce jitter and delays in Ethernet-based fronthaul source scheduling and preemption. In this section source scheduling algorithm is proposed and several considerations on its implementation are presented.

4.8.1 Algorithm for reducing UL jitter

The 802.1Qbv standard defines a framework for implementing traffic scheduling. However, an actual algorithm when to open and close the gates, is left outside of it and is up to the network configuration. The following algorithm is proposed for scheduling the traffic aiming at UL jitter reduction for fronthaul streams. It is optimized for symmetrical UL and DL, however, it can be used for any case when DL traffic is higher or equal than UL traffic.

The use cases are for one BBU per network and for:

1. 4G - CPRI traffic packetized into Ethernet frames. Many streams could share one link. The bit rate per each stream would be 2 Gbps instead of the original 2.5 Gbps when 8B/10B line coding is removed (Ethernet already implements error detecting coding by means of Frame check sequence (FCS)),

2. 5G - Fronthaul data from one of the new functional splits sent in the Ethernet packets. The bit rate per each stream would be between 150 Mbps and 2 Gbps, instead of a 2.5 Gbps CPRI. Optionally, for small cells the CPRI split can be used, as bit rates would be lower for small cells.

The BBU schedules the DL traffic to all the RRHs. The packets leave the output port of the BBU pool in order, which can be compared to virtual timeslots. For each packet received in DL, the RRH will send one UL packet. Therefore the UL packets will not collide if they are sent with correct timing advance fitting the virtual timeslots prepared for

DL. At the RRH bootup, control packets will need to be exchanged to measure the delay, in a similar way in which timing advance is implemented in GSM. Figure 4.20 describes the initial DL scheduling (left) and UL and DL packets fitting the virtual timeslots (right).

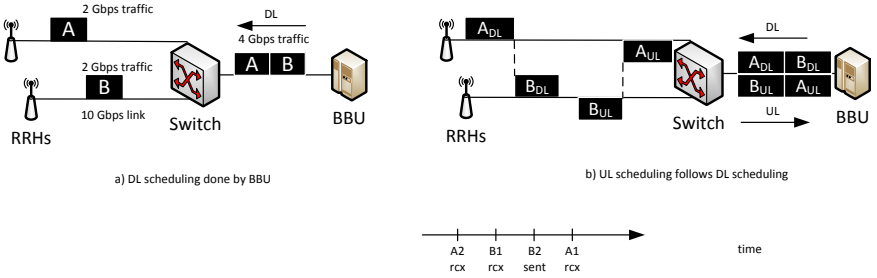


Figure 4.20: Source scheduling algorithm

In case more BBU pools are present in the network, a more generic solution will be needed. The methods used for wavelength assignment in wavelength routed networks [195] can be exploited. In case the network will be shared by other, non-fronthaul services it is recommended to add the preemption functionality.

4.8.2 Design considerations

The data need to be packetized into Ethernet frames. The choice of the optimal packet size is a trade off between a lower overhead for bigger sizes and smaller impact on lost or corrupted packet for smaller sizes. Figure 4.21 presents an overhead for various Maximum Transmission Unit (MTU) sizes starting from 128 B up to jumbo frame - 9000 B. Ethernet packet overhead was considered; on Layer 1: preamble (8 B), Interpacket gap (12 B), on Layer 2: frame header (14 B) and FCS (4 B) without an 802.1Q tag, in total 38 B, as well as 1904.3 overhead (10 B). For standard payload size of 1500 B the overhead is 3.2% and stays in the order of a few percent for higher sizes. Therefore 1500 B can be treated as a candidate packet size.

The algorithm described above can be implemented by means of control packets of the IEEE 1904.3. The framework given by IEEE 802.1Qbv can also be used.

The following procedure needs to be in place when RRH joins the network. Example implementation using IEEE 1904.3 RoE D0.1 is provided.

1. Stream setup

    a) BBU sends packet, notes the timestamp (T1)

    - High priority,
    - RoE pkt_type =ctrl (000000b), subtype=Access_1

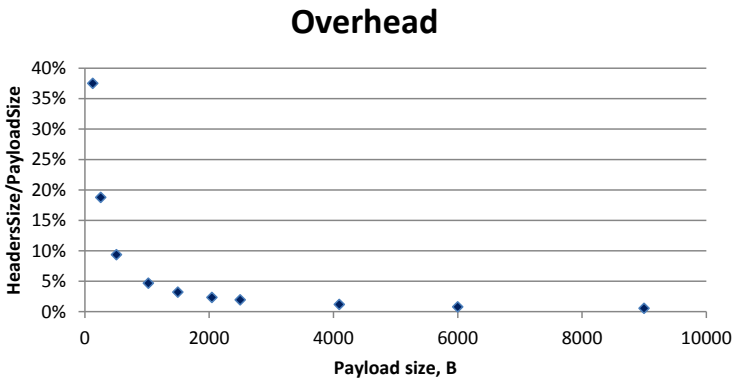


Figure 4.21: Ethernet L1 and L2 as well as 1904.3 overhead comparing to Ethernet frame payload size

b) Upon packet arrival RRH notes the time (T2), sends it back to BBU
   - Low priority,
   - RoE ctrl, subtype=Access_2

c) BBU computes delay Delay=T2-T1 informs RRH
   - Low priority,
   - RoE ctrl, subtype=Access_3

d) RRH acknowledges
   - Low priority,
   - RoE ctrl, subtype=Access_4

2. Data transmission

   a) BBU schedules traffic to the cell, sends it
      - RoE pkt_type =data(000001b - 000100b or 100100b),

   b) After receiving at least two packets m and m+1 at times Rcx(m) and Rcx(m+1), respectively, RRH measures receiving interval $RcxInt = Rcx(m+1) - Rcx(m)$

   c) Cell receives the traffic in time DLrcx, cell can send back the traffic in the earliest time in the future such that time $DLrcx - 2 \cdot Delay + n \cdot RcxInt$
      - RoE data

For the dependency between required RRH-BBU distance and allowed number of switches, the delay budget analysis presented in Section 4.7.1 applies.

4.9 Demonstrator of an Ethernet fronthaul

Previous sections described motivation for using packet-based fronthaul, associated challenges and solutions for meeting synchronization and delay requirements. This section presents a demonstrator of an Ethernet fronthaul that was built during this project and that is capable of transmitting one fronthaul stream. Future extensions are possible to enable network sharing between many streams and other applications.

In a joint laboratory work with Alcatel-Lucent Bell Labs France a demonstrator of an Ethernet-based fronthaul network was set up. A network consisting of 3 Ethernet switches controlled by an OpenDaylight SDN controller (default application) was built as presented in Figure 4.22. OpenFlow was used as a communication protocol. Two PCs, A and B, were running Linux, distribution Ubuntu version 14.04. It is a part of a C-RAN setup shown in Figure 4.23.

The delay measurements were performed using two methods:

1. Differential ping, method proposed by Henrik Christiansen,
2. DPDK application.

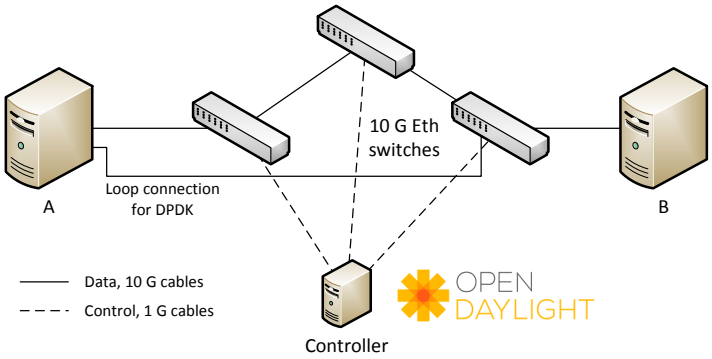


Figure 4.22: Demonstrator of Ethernet fronthaul network

4.9.1 Delay measurements using differential ping method

A ping test was performed for 1-3 switches for packet of different sizes (0 - 8972 B). The sample results are shown on Figure 4.24. The delay consists of the following factors, as introduced in Section 4.7:

$$Delay = PropDel + SwitchDel + TransDel(B) + queuing \tag{4.10}$$

Here the distance was within few meters, therefore:

$$PropDel \approx ns \tag{4.11}$$

$$PingRTT = 2 \cdot Delay + TimeInPCs \tag{4.12}$$

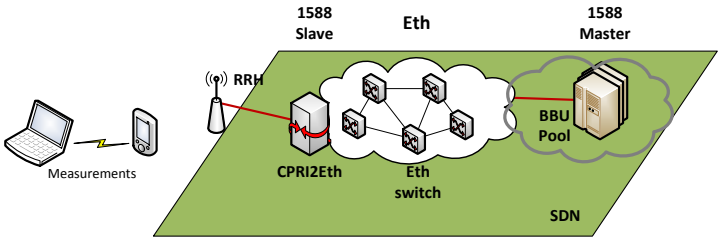


Figure 4.23: Ethernet-based C-RAN fronthaul - laboratory setup

Table 4.5: Analysis of a ping delay

# switches	minRTT (0B), μs	Diff to #-1 (μs)	Jitter, μs	Avg jitter, μs
1	53.4	-	18-210	120
2	58.2	4.8	16-203 and 576	112
3	62.5	4.3	22-384 and 575	120

$TimeInPCs$ is the one needed for the kernel to process packets in each of the PCs. In order to calculate the switching delay ($SwitchDel$), the values where the minimum RTT crosses the y axes were noted down, corresponding to the delay where transmission delay depending on packet size ($TransDel(B)$) is zero and queuing delay is also zero - e.g. $PingRTT1switch(0)$ for one switch. That leads to:

$$PingRTT1switch(0) = 2 \cdot SwitchDelay + TimeInPCs \quad (4.13)$$

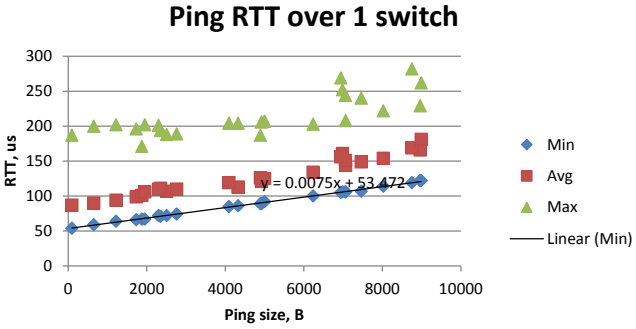
$$PingRTT2switches(0) = 4 \cdot SwitchDelay + TimeInPCs \quad (4.14)$$

$$PingRTT3switches(0) = 6 \cdot SwitchDelay + TimeInPCs \quad (4.15)$$

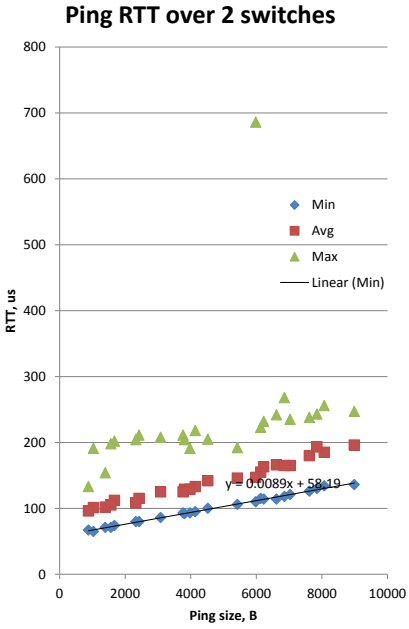
By subtracting the $PingRTTNswitch(0)$ values (and dividing the difference by 2), as shown in Table 4.5, an average switch delay is 2.3 μs.

4.9.2 Delay measurements using DPDK application

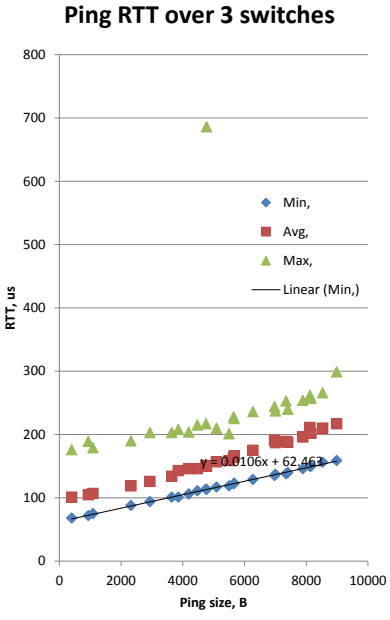
In order to test the delay more accurately a measurement setup was prepared using a Data Plane Development Kit (DPDK). DPDK is a set of libraries and drivers for fast packet processing that runs on an Ethernet card. In that way the $TimeInPCs$ is reduced to minimum. The measurements were performed between 2 ports of the same PC (A in Figure 4.22), in order to measure time with a common clock. Traffic from 1 Gbps to 9 Gbps was sent over 1, 2 and 3 switches. The results are summarized in Table 4.6. Average $SwitchDelay$ is 2.1 μs, thereby the processing time is PC $TimeInPCs$ can be estimated to 0.4 μs.



(a)



(b)



(c)

Figure 4.24: Ping RTT over 1 - 3 switches

Table 4.6: Delay measurements over a dedicated line using DPDK

# switches	Avg Delay, (μs)	Max Jitter, (μs)
1	2.5	0.5
2	4.6	0.6
3	6.7	0.6

4.9.3 Summary and discussion

A demonstrator capable of transmitting a single CPRI stream over Ethernet fulfilling the delay requirements was built. Measured delays are as expected for a store-and-forward switch, in the order of μs. The measured value for the Differential ping, which is a simple method, was 2.3 μs, while for DPDK, which required a DPDK application and more complicated configuration, was 2.1 μs. Therefore a Differential ping proved to be a sufficiently accurate method. However, measurements using DPDK were more accurate and served as a useful verification. For Differential ping method $TimeInPCs$ was over 50 μs, while for DPDK 0.4 μs. This huge difference is due to the fact that higher protocol layers were processed and more PC components were involved to measure ping than using only an Ethernet card for DPDK application.

4.10 Future directions

The new functional split between RRH and BBU will result in variable bit rate data, that can be transported in a similar way to backhaul traffic. Therefore the work on both are converged under the term *Xhaul* [196], [197]. X-haul refers to the network segment optimized for transporting any type of data, especially packet-based data, as opposed to IQ samples carried by traditional fronthaul. It should not be confused with *midhaul* which is defined by Metro Ethernet Forum (MEF) [174] as network between base stations, typically between one macro base station and several small cell sites. It is illustrated in Figure 4.25. Fronthaul is an intra-base station connection, while midhaul is an inter-base station connection.

Current delay requirements result from HARQ process timeout of 8 ms. For 5G networks it is considered to relax this requirement. Also, if out of this 8ms both UE and base station would need less than 3ms the budget for fronthaul could be extended.

4.11 Summary of standardization activities

Fronthaul evolution leads to different fronthaul splits for different deployment scenarios most likely resulting in packet-based transport on a fronthaul network. To enable that the following standardization activities are in place:

- IEEE NGFI is under preparation to define new functional splits.

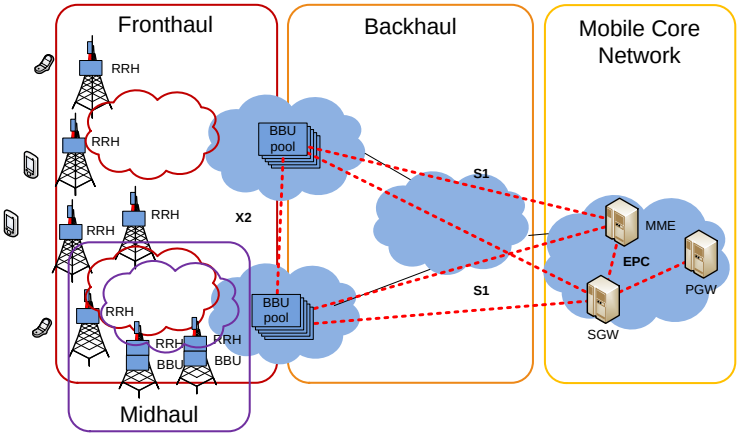


Figure 4.25: Fronthaul, backhaul and midhaul

- IEEE 1904.3 "Standard for Radio Over Ethernet Encapsulations and Mappings" that aims to define encapsulation of fronthaul data, independent on the functional split, to Ethernet frames.

- IEEE 802.1Qbv "Bridges and Bridged Networks–Amendment: Enhancements for Scheduled Traffic" standardizes traffic scheduling in order to reduce jitter of fronthaul streams in Ethernet switches by ensuring that, at specific times, only one traffic stream has access to the network.

- IEEE 802.1Qbu "Bridges and Bridged Networks–Amendment: Frame Preemption" standardizes preemption which can suspend the transmission of the lower priority traffic when fronthaul traffic needs to be transmitted.

- IEEE 802.1CM "Time-Sensitive Networking for Fronthaul" profile that select features, options, configurations, defaults, protocols and procedures defined in IEEE 802 standard for switches and stations for a fronthaul application.

- IEEE 1588 "Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems" v3 is under preparation which aims, among others to improve accuracy to 100s ps and enhance security. The expected completion date of v3 of the standard is 31 December 2017 [7].

IEEE 802.1Qbv, IEEE 802.1Qbu and IEEE 802.1CM are part of TSN Task Group.

4.12 Summary

In order to lower the costs and improve the flexibility of C-RAN deployments, existing packet-based, e.g. Ethernet, networks can be reused for fronthaul.

Sections above analyze requirements on bit rates, delay, phase and frequency synchronization for fronthaul with current and new functional splits. Architecture sufficient for fulfilling them was derived, namely: CPRI/OBSAI over OTN and CPRI/new functional split data over Ethernet. It has been proved that OTN can support existing deployments with CPRI/OBSAI. Factors that are challenging for achieving synchronization in packet-based C-RAN fronthaul were analyzed. A feasibility study was presented showing the performance for frequency and phase synchronization using 1588 in Ethernet networks under various inaccuracies that can be present in the network. Apart from possible queuing delays, the one that has the highest impact is the timestamping error associated with the way timestamps are generated in Ethernet switches. Whether this performance will meet the requirements of future mobile network depends on PLL and local oscillator implementation, based on 1588 feedback to clock offset and drift. Moreover, an Ethernet network that is ready to be integrated in the fronthaul part of a C-RAN demo and which fulfills the delay requirements was built. The delay measurements were performed, which allowed to obtain a better understanding of the delays encountered in an Ethernet network. To address the delay requirements in shared networks source scheduling and preemption are investigated. A source scheduling algorithm is proposed, which is optimized for symmetrical fronthaul traffic, however can also be applied in cases when downlink traffic exceeds uplink traffic.

Figure 4.26 presents a protocol stack enabling various functionalities for fronthaul over Ethernet transport. Devices at the cell site can either be legacy RRHs running on CPRI and connected to a CPRI2Eth gateway or native Ethernet RRHs. 1588 assures synchronization between a master clock located in a BBU pool and a slave clock located at the cell site. It is beneficial to allow on path support of 1588 for better compensation of queuing related delays, however, in principle it is not mandatory. It is recommended to use TCs for a better network delay compensation. If an Ethernet network consists of unknown types of switches, non-TSN enabled, it is recommended to enable source scheduling at the network edges to control the delay. Moreover, if it is desired to share the network with other services, preemption is a recommended feature for Ethernet switches.

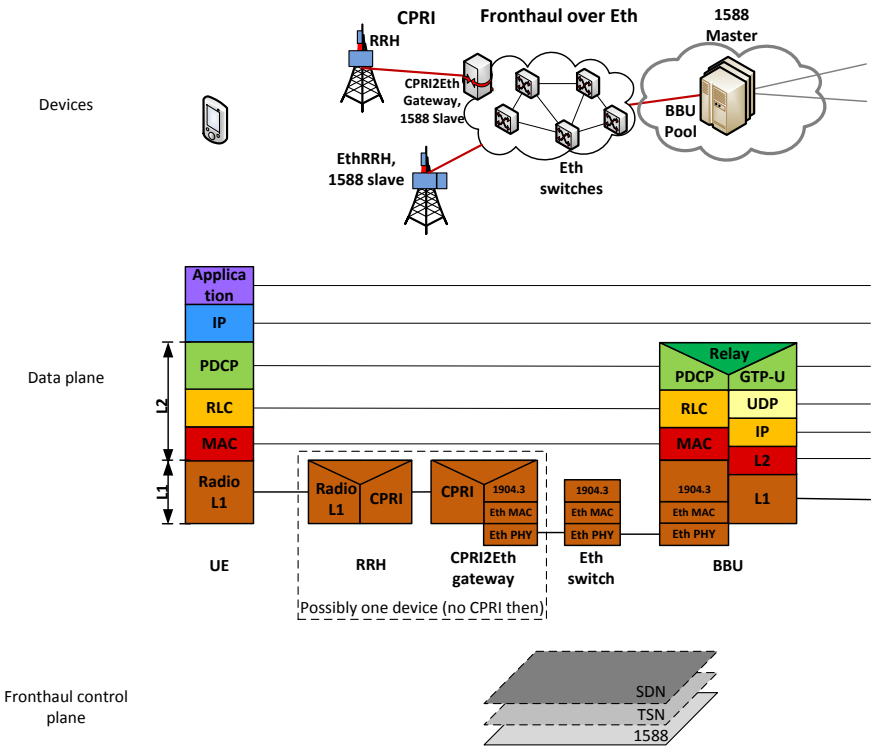


Figure 4.26: Proposed architecture for Fronthaul over Ethernet. Dashed lines highlight optional functionality.

CHAPTER 5

Conclusions and outlook

In order to satisfy the ever growing need for capacity in mobile networks, and, at the same time, to create cost and energy efficient solutions, new and disruptive ideas are needed. With its performance gains and cost benefits, C-RAN proves to be a major technological foundation for 5G mobile networks. By applying the concept of NFV, C-RAN mobile networks are following the IT paradigm towards virtualization and cloudification. Decoupling hardware from software enables software to orchestrate various components, including BBU pool and network resources. Such a flexible, automated, and self organized network allows for various optimizations.

C-RAN builds on the base station architecture evolution from a traditional base station, via a base station with RRH to a centralized and virtualized one. In C-RAN base station functionalities are split between cell locations and the centralized pool. It is a challenge to find an optimal splitting point as well as assuring efficient interconnectivity between the parts. This dissertation summarizes the efforts to analyze this architecture, evaluate its benefits towards energy and cost savings as well as investigates a flexible fronthaul architecture. To conclude, C-RAN is a candidate architecture worth considering for 5G deployments to address its performance needs and optimize deployment costs. Centralization and virtualization indeed offer cost savings. The point where the functionality needs to be split between the cell site and the centralized location needs to be chosen for a particular deployment, and a variable bit rate splits are foreseen. Ethernet-based fronthaul is not straightforward to implement. However, with the discussed techniques it has the potential to meet mobile networks' requirements.

A comprehensive overview of C-RAN is presented. Details of this architecture are provided, along with its benefits and technical challenges. Due to sharing of baseband resources, C-RAN adapts well to traffic fluctuations among cells. Capacity can be scaled more easily, energy and cost of baseband units pool deployment and operation can be lowered. Moreover, cooperative techniques are enhanced, increasing the overall cell throughput, while dealing efficiently with interference. However, requirements of high capacity and low delay are put on the fronthaul network, and development of virtualization techniques is needed. Answers to those challenges follow in terms of: analysis of possible transport network mediums and techniques, the needed RRH and BBU development as well as an overview of virtualization techniques. Likely deployment scenarios are presented together with a broad overview of industry and academic work on developing and boosting C-RAN up to the beginning of 2014. Last, but not least, an overview of future directions for C-RAN is provided with a strong focus on defining a new functional

split between BBU and RRH in order to enable flexible fronthaul and lower data bit rate in this network segment.

One of the main advantages of C-RAN, multiplexing gains, is thoroughly analyzed. Various sources of multiplexing gains have been identified and quantified for traffic-dependent resources, namely: users changing location daily between e.g. work and home - the so called tidal effect, traffic burstiness, as well as different functional splits. For traditional C-RAN deployments, with a functional split between baseband and radio functionalities, a multiplexing gain can be achieved on baseband units. However, when the functional split allows for variable bit rate, multiplexing gains can also be exploited on fronthaul links. The latter is an important motivation and guideline for designing the functional split not only to lower the bit rate, but also costs of the fronthaul network. For the analyzed data sets, the multiplexing gain value reaches six, in deployments where various traffic types are mixed (bursty, e.g. web browsing and constant bit rate, e.g. video streaming) and cells from various areas (e.g. office and residential) are connected to the same BBU pool.

In order to further optimize the cost of C-RAN deployments, the possibility of reusing existing Ethernet networks has been exploited. Such an architecture is especially optimal for functional splits resulting in variable bit rate traffic in the fronthaul. Packet-based networks enable multiplexing gains and flexible, multipoint-to-multipoint connectivity between cell sites and centralized locations. However, assuring synchronization and meeting stringent delay requirements is a challenge. Mechanisms for delivering a reference clock to cell sites have been analyzed, and an architecture employing IEEE 1588, also known as PTP, has been evaluated as a candidate technology for C-RAN. For the tested CPRI-like scenario, the proposed filtering gave sufficient accuracy to fulfill the requirements of mobile networks - in the order of nanoseconds. Regarding the delay requirements, the sources of delays have been identified and quantified. The non-deterministic queuing delay is challenging because of its possibly variable value, however, it can be addressed with: 1) source scheduling, especially at the edge of the network, and 2) preemption in the switches. A source scheduling algorithm has been proposed to address the jitter and delay constraints. It is optimized for cases where downlink traffic equals uplink traffic, but can also be used when downlink traffic exceeds uplink. Moreover, a demonstrator of an SDN controlled Ethernet-based fronthaul has been prepared.

This thesis identifies several possible directions for future work, both to further improve its findings and to explore a wider perspective.

5.1 Future research

Five years until the standardization of 5G networks is a busy time for the mobile networks ecosystem. Several directions are investigated to enable further capacity growth and support the increasing number of devices, especially for IoT. This section attempts to discuss possible directions for fronthaul networks development.

The traditional base station **functional split** between baseband and radio functionalities was fine for short scale deployments, e.g., between rooftop and basement. However, with C-RAN, bringing fronthaul into metropolitan scale with ever-growing capacity needs, more disruptive solutions are needed. When this project was started, the main focus to address this issue was on compression. Now, towards the end of the project, new functional spits are being discussed: load-dependent and multipoint-to-multipoint. The user centric/cell centric split is a very promising one, as it brings the fronthaul data rate almost to the level offered for users, creating backhaul-like traffic, still leaving many functions centralized. **Fronthaul networks are receiving high interest from the standardization bodies**, in terms of both defining a functional split and on meeting synchronization and delay requirements. The main examples are IEEE NGFI, IEEE 1904.3, and 802.1CM - TSN profile for fronthaul.

This project concentrated on **studying multiplexing gains** on resources dependent on user traffic. As there are parts of the base station that need to be on to provide coverage, even when no users are active, it would be beneficial to study overall multiplexing gains including those modules, too. Moreover, another method could be used to quantify the multiplexing gains, not applying traffic averaging, but possibly quantifying computational resources in terms of e.g., the number of operations per second.

In this project, a CPRI-like scenario, including one data stream was analyzed for **synchronization accuracy**. Future work could investigate how the varying network load for several data streams from a variable-bit rate splits affects synchronization accuracy.

The proposed source scheduling algorithm does not cover scenarios with uplink traffic exceeding downlink. Moreover, it is assumed that only one BBU pool is present in the network. A **more generalized algorithm for source scheduling**, with multiple BBU pools and possibly other services present in the network, is of interest for 5G.

One of the main constraints for fronthaul is the **delay requirement** coming from the HARQ process. With the current design, not meeting this requirement results in retransmission, thereby lowering application data throughput. Therefore more studies are needed on whether HARQ is needed in 5G, if the timer could be extended, or what could be the delay budget share between base station and UE.

Moreover, substituting fibers with a **wireless fronthaul** is an important research and development topic. Exploring new frequency bands, together with increasing spectral efficiency of transport via this medium can address growing capacity needs.

Last but not least, solutions for hardware sharing are of high interest. Virtualization enables sharing e.g. baseband resources, however methods for sharing other elements, like fronthaul or cell site equipment, are important, too. The topic of network sharing is connected with **hybrid fronthaul and backhaul optimization**. Future mobile networks will most likely consist of standalone base stations as well as base stations aggregated in a C-RAN. Joint capacity and control plane optimization for midhaul networks will enable more efficient usage of resources, thereby lowering network deployment and operation costs.

Appendices

APPENDIX A

OTN-based fronthaul

The following sections on OTN demonstrator are published in [86].

A.1 OTN solution context

When the customer/mobile network operator owns little fiber or when the cost of leasing fiber is high in both 1st and 2nd mile, the C-RAN architecture presented in Figure A.1 is beneficial. Baseband pool (BBU Pool) is located on OTN ring. CPRI/OBSAI is carried between BBU Pool and RRH over OTN. CPRI/OBSAI can be mapped to OTN containers using the OTN mapper from Altera.

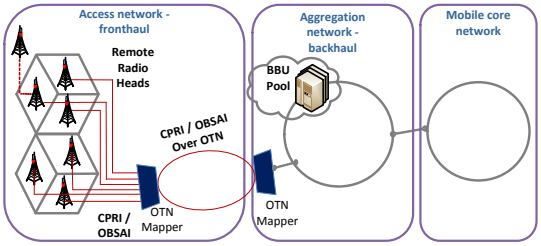


Figure A.1: C-RAN architecture where OTN is used to transport fronthaul streams

A.2 Overview

We benchmarked the CPRI/OBSAI over OTN transport performance against a reference setup shown in Figure A.2. BSE sends IQ data, here over CPRI protocol to RRH. A signal analyzer is used to measure EVM and Frequency Error of the transmitted signal from the RRH antenna port.

The actual measurement setup is shown in Figure A.3. We introduced the Altera TPO124/125 OTN multiplexer that maps CPRI client signals to OTN containers and back from OTN containers to CPRI. We measured data EVM and Frequency Error and compare it to the one achieved with the setup presented in Figure A.2. The detailed overview of the system is presented in Figure A.4.

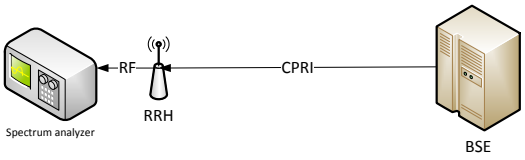


Figure A.2: Reference setup for CPRI over OTN testing

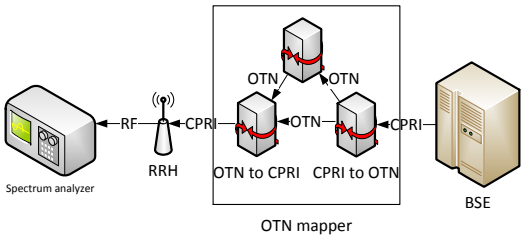


Figure A.3: CPRI over OTN mapping measurement setup

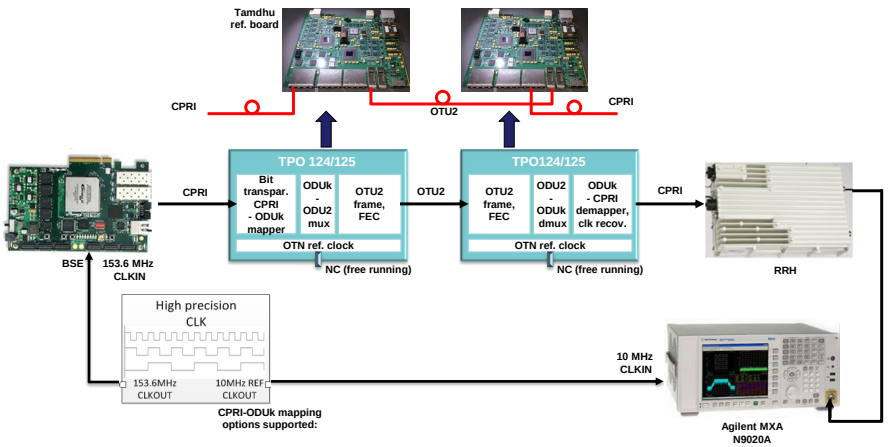


Figure A.4: Detailed measurement setup [86]

Four separate sets of measurements were conducted, focusing on CPRI and OBSAI, using TPO124 and TPO125 mappers as summed up in Table A.1. Parameters are summarized in Table A.2.

Table A.1: Measurement scenarios

Client \ OTN Mapper	TPO124	TPO125
CPRI-3 2.4576 Gbps	Scenario 1	Scenario 2
OBSAI 3.072 Gbps	Scenario 3	Scenario 4

A.3 Results

Table A.3 summarizes the results for CPRI and OBSAI protocols. For three different modulations: QPSK, 16 QAM and 64 QAM measurements were taken without OTN device, as presented in Figure A.2, and then with TPO124 and TPO125 as OTN mappers, as presented in Figure A.3. We present maximum observed EVM and frequency error for each modulation, for each scenario. Table A.3 presents worst case observations over a 1-minute interval.

Looking at the performance of transmission of CPRI over OTN, we can see that OTN transmission caused negligible EVM comparing to reference scenario. The frequency error increased; however, it stays within the requirements. The performance of OBSAI over OTN transmission should be compared to its reference scenario. The conclusions are similar: negligible EVM increase, frequency error stays within requirements.

Figure A.5 shows an example of results for 64 QAM modulated signal transmitted with OBSAI protocol over OTN with TPO125 device. Upper left figure shows the modulation constellation. In the upper right figure the frequency error is displayed for each of the 20 slots of a 10 ms LTE frame. Lower right figure shows summary of measurements with EVM and averaged Frequency Error.

A.4 Conclusion

OTN is a posible optical transport solution for IQ transport between RRH and BBU Pool when mobile network operator has a cost-efficient access to legacy OTN network, which can be reused for C-RAN. We present a proof of concept of transmitting radio interface protocols over OTN enabling to exploit benefits of C-RAN. Tested solution introduces negligible EVM increase and small frequency error. It is fully compliant with 3GPP requirements for LTE-Advanced. Future work could include integration of setup with higher CPRI/OBSAI bit rates up to 10 Gbps and introducing verification of deterministic delay measurements.

Table A.2: Setup specifications

Base Station	Base Station Emulator	Multimode
	External clock	153.6 MHz
Optical transport	Fiber	
	SFP	Multimode, 850nm, Finisar FTLF8524P2xNy
OTN	TPO124	Talisker Reference board with TPO124 client-to-OTU2 mapper
	TPO125	Tamdhu Reference board with TPO125 client-to-OTU2 mapper
	Optical fiber loopback on OTU2 port	
RRH for CPRI	Model	RRH 700 MHz, FDD, 2x67 W
	Carrier center frequency	737 MHz
RRH for OBSAI	Model	RRH 850 MHz, FDD, 2x40 W
	Carrier center frequency	880 MHz
Client: CPRI-3 2.4576 Gbps	Mapping	GMP/ODU1
	Multiplexing	AMP ODU1/ODU2
	Line	OTU2
Client: OBSAI 3.072 Gbps	Mapping	BMP/ODUflex (as per CPRI-4)
	Multiplexing	GMP ODUflex/ODU2
	Line	OTU2
Test signals [81]	LTE, 10MHz, 10 ms	
	QPSK	ETM 3.3
	16 QAM	ETM 3.2
	64 QAM	ETM 3.1

Table A.3: Measurements results summary

Modulation	OTN device	CPRI		OBSAI		Requirements	
		EVM, %	Frequency Error, ppm	EVM, %	Frequency Error, ppm	EVM, % [81]	Frequency Error, ppm
QPSK	-	5.5	0.003	9.9	0.005	<17.5%	<0.05 ppm [81]
	TPO124	5.7	0.026	9.9	0.034		
	TPO125	5.7	0.015	-	-		
16 QAM	-	5.5	0.003	6.9	0.007	<12.5%	
	TPO124	5.6	0.023	7.2	0.028		
	TPO125	5.6	0.016	-	-		
64 QAM	-	5.5	0.003	4.4	0.005	<8%	
	TPO124	5.7	0.027	5.1	0.028		
	TPO125	5.7	0.018	4.5	0.034		

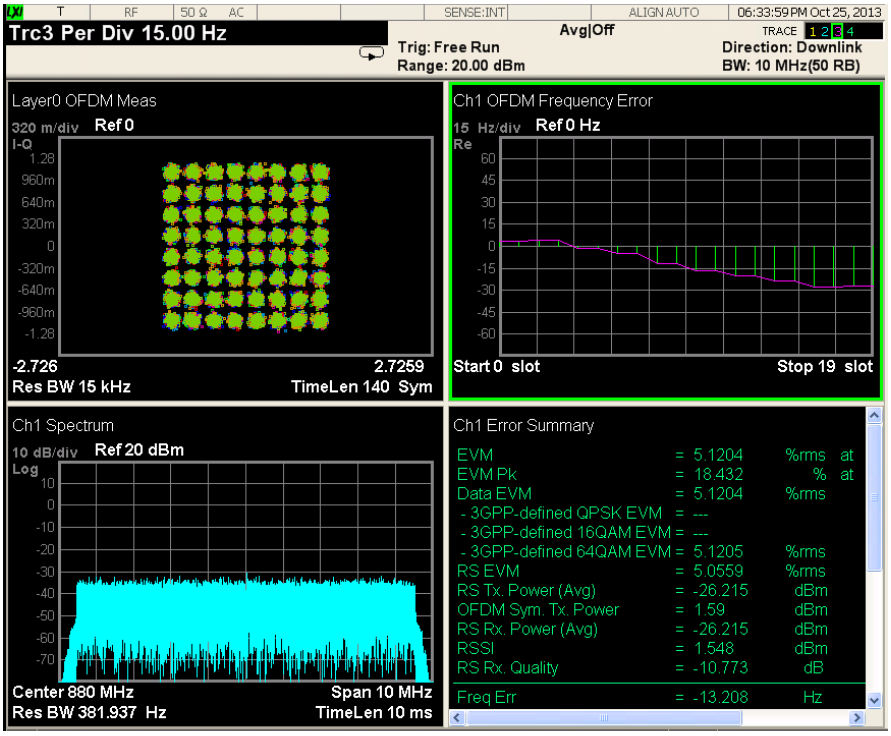


Figure A.5: Results 64 QAM with OBSAI Using TPO125 Device

Bibliography

[1] *Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017*. Tech. rep. Cisco, February 2013 (cit. on pp. 2, 64).

[2] *Mobility Report*. Tech. rep. Ericsson, November 2015 (cit. on pp. 2, 73).

[3] J. Gozalvez. "Tentative 3GPP Timeline for 5G [Mobile Radio]". In: *Vehicular Technology Magazine, IEEE* 10.3 (Sept. 2015), pp. 12–18. ISSN: 1556-6072. DOI: 10.1109/MVT.2015.2453573 (cit. on p. 2).

[4] *5G White Paper*. Tech. rep. NGMN Alliance, February 2015 (cit. on pp. 2, 3).

[5] *Mobility Report*. Tech. rep. Ericsson, November 2011 (cit. on p. 2).

[6] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao. "5G on the Horizon: Key Challenges for the Radio-Access Network". In: *Vehicular Technology Magazine, IEEE* 8.3 (Sept. 2013), pp. 47–53. ISSN: 1556-6072. DOI: 10.1109/MVT.2013.2269187 (cit. on p. 3).

[7] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan. "Toward Green and Soft: a 5G Perspective". In: *Communications Magazine, IEEE* 52.2 (Feb. 2014), pp. 66–73. ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6736745 (cit. on p. 3).

[8] L. B. Le1, V. Lau, E. Jorswieck, N.-D. Dao, A. Haghighat, D. I. Kim, and T. Le-Ngoc. "Enabling 5G mobile wireless technologies". In: *EURASIP Journal onWireless Communications and Networking* (Dec. 2015). DOI: DOI10.1186/s13638-015-0452-9 (cit. on p. 3).

[9] M. Peng, Y. Li, Z. Zhao, and C. Wang. "System architecture and key technologies for 5G heterogeneous cloud radio access networks". In: *Network, IEEE* 29.2 (Mar. 2015), pp. 6–14. ISSN: 0890-8044. DOI: 10.1109/MNET.2015.7064897 (cit. on p. 3).

[10] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann. "Cloud RAN for Mobile Networks - a Technology Overview". In: *IEEE Communications Surveys & Tutorials, IEEE* 17.1 (Firstquarter 2015). ©2015 IEEE. Reprinted, with permission, pp. 405–426. ISSN: 1553-877X. DOI: 10.1109/COMST.2014.2355255 (cit. on pp. 4, 9).

[11] A. Checko, H. Christiansen, and M. S. Berger. “Evaluation of energy and cost savings in mobile Cloud-RAN”. In: *Proceedings of OPNETWORK Conference*. 2013 (cit. on pp. 4, 15, 42, 50, 62, 76).

[12] A. Checko, H. Holm, and H. Christiansen. “Optimizing small cell deployment by the use of C-RANs”. In: *European Wireless 2014; 20th European Wireless Conference; Proceedings of*. ©2014 VDE. Reprinted, with permission. May 2014, pp. 1–6 (cit. on pp. 4, 15, 16, 37, 42, 50, 67, 68, 76).

[13] A. Checko[1st], A. P. Avramova[1st], H. L. Christiansen, and M. S. Berger. “Evaluating C-RAN fronthaul functional splits in terms of network level energy and cost savings”. In: *accepted to IEEE Journal Of Communications And Networks* () (cit. on pp. 4, 16, 50, 64, 75, 76).

[14] A. Checko, L. Ellegaard, and M. Berger. “Capacity planning for Carrier Ethernet LTE backhaul networks”. In: *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. Apr. 2012, pp. 2741–2745. DOI: 10.1109/WCNC.2012.6214266 (cit. on pp. 4, 76).

[15] A. Checko, A. Juul, H. Christiansen, and M. Berger. “Synchronization challenges in packet-based Cloud-RAN fronthaul for mobile networks”. In: *Communication Workshop (ICCW), 2015 IEEE International Conference on*. ©2015 IEEE. Reprinted, with permission. June 2015, pp. 2721–2726. DOI: 10.1109/ICCW.2015.7247590 (cit. on pp. 4, 83).

[16] I. Hwang, B. Song, and S. Soliman. “A Holistic View on Hyper-Dense Heterogeneous and Small Cell Networks”. In: *Communications Magazine, IEEE* 51.6 (2013). ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6525591 (cit. on pp. 7, 8).

[17] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer. “Shifting the MIMO Paradigm”. In: *Signal Processing Magazine, IEEE* 24.5 (2007), pp. 36–46. ISSN: 1053-5888. DOI: 10.1109/MSP.2007.904815 (cit. on p. 7).

[18] J. Hoydis, S. ten Brink, and M. Debbah. “Massive MIMO: How many antennas do we need?” In: *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. 2011, pp. 545–550. DOI: 10.1109/Allerton.2011.6120214 (cit. on p. 7).

[19] H. Guan, T. Kolding, and P. Merz. *Discovery of Cloud-RAN*. Tech. rep. Nokia Siemens Networks, April 2010 (cit. on pp. 7, 8, 17, 42).

[20] China Mobile Research Institute. *C-RAN The Road Towards Green RAN*. White Paper, Version 2.5, October 2011 (cit. on pp. 7, 8, 14–17, 20, 21, 24, 25, 28–30, 33, 34, 41, 42, 49, 56, 62, 64, 68, 84, 85).

[21] Y. Lin, L. Shao, Z. Zhu, Q. Wang, and R. K. Sabhikhi. "Wireless network cloud: Architecture and system requirements". In: *IBM Journal of Research and Development* 54.1 (Jan. 2010), 4:1–4:12. ISSN: 0018-8646. DOI: 10.1147/JRD.2009.2037680 (cit. on pp. 7, 8, 12, 42).

[22] J. Segel. *lightRadio Portfolio: White Paper 3*. Tech. rep. Alcatel-Lucent Bell Labs, 2011 (cit. on pp. 8, 21, 25, 29, 30, 32, 33, 42).

[23] Huawei. *Cloud RAN Introduction. The 4th CJK International Workshop –Technology Evolution and Spectrum*. September 2011 (cit. on pp. 8, 26, 42).

[24] *ZTE Green Technology Innovations White Paper*. Tech. rep. ZTE, 2011 (cit. on pp. 8, 17, 42).

[25] *Intel Heterogenous Network Solution Brief*. Tech. rep. 2011 (cit. on pp. 8, 33, 34, 41, 42).

[26] T. Flanagan. *Creating cloud base stations with TI's KeyStone multicore architecture*. Tech. rep. Texas Instruments, October 2011 (cit. on pp. 8, 22, 34, 42).

[27] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan. "Toward Green and Soft: a 5G Perspective". In: *Communications Magazine, IEEE* 52.2 (Feb. 2014), pp. 66–73. ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6736745 (cit. on p. 8).

[28] *lightRadio Network. Alcatel-Lucent*. [cited: January 2016]. URL: https://www.alcatel-lucent.com/solutions/lightradio (cit. on p. 8).

[29] W. Liu, S. Han, C. Yang, and C. Sun. "Massive MIMO or Small Cell Network: Who is More Energy Efficient?" In: *Wireless Communications and Networking Conference Workshops (WCNCW), 2013 IEEE*. 2013, pp. 24–29. DOI: 10.1109/WCNCW.2013.6533309 (cit. on p. 8).

[30] J. Madden. *Cloud RAN or Small Cells?* Tech. rep. Mobile Experts, April 2013 (cit. on pp. 8, 42).

[31] G. Kardaras and C. Lanzani. "Advanced Multimode Radio for Wireless and Mobile Broadband Communication". In: *European Wireless Technology Conference, 2009. EuWIT 2009*, pp. 132–135 (cit. on p. 10).

[32] E. Dahlman, S. Parkvall, J. Skold, and P. Beming. *3G Evolution: HSPA and LTE for Mobile Broadband*. 3G Evolution. Elsevier Science, 2010. ISBN: 9780080923192 (cit. on p. 10).

[33] *Common Public Radio Interface (CPRI); Interface Specification V6.0*. Aug. 2013 (cit. on pp. 12, 21, 31, 87).

[34] *Open Base Station Architecture Initiative (OBSAI) BTS System Reference Document Version 2.0*. 2006 (cit. on p. 12).

[35] *ETSI GS ORI 002-1 V1.1.1 (2011-10). Open Radio equipment Interface (ORI); ORI Interface Specification; Part 1: Low Layers (Release 1)* (cit. on p. 12).

[36] *ETSI GS ORI 002-2 V1.1.1 (2012-08). Open Radio equipment Interface (ORI); ORI Interface Specification; Part 2: Control and Management (Release 1)* (cit. on p. 12).

[37] S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao. "Distributed wireless communication system: a new architecture for future public wireless access". In: *Communications Magazine, IEEE* 41.3 (Mar. 2003), pp. 108–113. ISSN: 0163-6804. DOI: 10.1109/MCOM.2003.1186553 (cit. on p. 12).

[38] F. Anger. *Smart Mobile Broadband. In proceedings of RAN Evolution to the Cloud Workshop*. June 2013 (cit. on p. 14).

[39] G. Brown. *Converging Telecom & IT in the LTE RAN*. Tech. rep. Samsung, Feb 2013 (cit. on p. 14).

[40] S. Namba, T. Matsunaka, T. Warabino, S. Kaneko, and Y. Kishi. "Colony-RAN architecture for future cellular network". In: *Future Network Mobile Summit (FutureNetw), 2012*. July 2012, pp. 1–8 (cit. on pp. 14, 15, 22, 42, 78).

[41] T. Werthmann, H. Grob-Lipski, and M. Proebster. "Multiplexing gains achieved in pools of baseband computation units in 4G cellular networks". In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. Sept. 2013, pp. 3328–3333. DOI: 10.1109/PIMRC.2013.6666722 (cit. on pp. 14, 42, 55, 78).

[42] S. Bhaumik, S. P. Chandrabose, M. K. Jataprolu, A. Muralidhar, V. Srinivasan, G. Kumar, P. Polakos, and T. Woo. "CloudIQ: A framework for processing base stations in a data center". In: *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM* (2012), pp. 125–136 (cit. on pp. 15, 33, 42).

[43] M. Madhavan, P. Gupta, and M. Chetlur. "Quantifying multiplexing gains in a Wireless Network Cloud". In: *Communications (ICC), 2012 IEEE International Conference on*. 2012, pp. 3212–3216. DOI: 10.1109/ICC.2012.6364658 (cit. on pp. 15, 42).

[44] J. Liu, S. Zhou, J. Gong, Z. Niu, and S. Xu. "On the statistical multiplexing gain of virtual base station pools". In: *Global Communications Conference (GLOBECOM), 2014 IEEE*. Dec. 2014, pp. 2283–2288. DOI: 10.1109/GLOCOM.2014.7037148 (cit. on p. 15).

[45] A. Avramova, H. Christiansen, and V. Iversen. "Cell Deployment Optimization for Cloud Radio Access Networks using Teletraffic Theory". In: *The Eleventh Advanced International Conference on Telecommunications, AICT 2015*. June 2015 (cit. on pp. 15, 67, 68).

[46] *White Paper of Next Generation Fronthaul Interface*. Tech. rep. China Mobile Research Institute, Alcatel-Lucent, Nokia Networks, ZTE Corporation, Broadcom Corporation, Intel China Research Center, October 2015 (cit. on pp. 16, 44, 51, 52, 84, 85, 88).

[47] C. Liu, K. Sundaresan, M. Jiang, S. Rangarajan, and G.-K. Chang. "The case for reconfigurable backhaul in Cloud-RAN based small cell networks". In: *INFOCOM, 2013 Proceedings IEEE*. Apr. 2013, pp. 1124–1132. DOI: 10.1109/INFCOM.2013.6566903 (cit. on pp. 16, 41, 42).

[48] H. Jinling. "TD-SCDMA/TD-LTE evolution - Go Green". In: *Communication Systems (ICCS), 2010 IEEE International Conference on*. 2010, pp. 301–305. DOI: 10.1109/ICCS.2010.5686439 (cit. on pp. 16, 17, 42).

[49] C. Chen. *C-RAN: the Road Towards Green Radio Access Network. Presentation*. August 2012 (cit. on pp. 17, 41, 42).

[50] *C-RAN - Road Towards Green Radio Access Network. Centralized baseband, Collaborative radio, and real-time Cloud computing RAN. Presentation*. EXPO 2010 (cit. on pp. 17, 42).

[51] J. Acharya, L. Gao, and S. Gaur. *Heterogeneous Networks in LTE-Advanced*. Wiley, 2014. ISBN: 9781118693957. URL: https://books.google.dk/books?id=RiDnAgAAQBAJ (cit. on p. 17).

[52] P. Marsch and G. Fettweis. *"Coordinated Multi-Point in Mobile Communications: From Theory to Practice"*. Cambridge University Press, 2011. ISBN: 9781107004115 (cit. on p. 17).

[53] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel. "Coordinated multipoint: Concepts, performance, and field trial results". In: *Communications Magazine, IEEE* 49.2 (2011), pp. 102–111. ISSN: 0163-6804. DOI: 10.1109/MCOM.2011.5706317 (cit. on pp. 17, 22, 42).

[54] H. Holma and A. Toskala. *"LTE-Advanced: 3GPP Solution for IMT-Advanced"*. John Wiley and Sons, Ltd, 2012. ISBN: 9781119974055 (cit. on pp. 17, 20, 23, 42).

[55] Y. Huiyu, Z. Naizheng, Y. Yuyu, and P. Skov. "Performance Evaluation of Coordinated Multipoint Reception in CRAN Under LTE-Advanced uplink". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 778–783. DOI: 10.1109/ChinaCom.2012.6417589 (cit. on pp. 17, 42).

[56] L. Li, J. Liu, K. Xiong, and P. Butovitsch. "Field test of uplink CoMP joint processing with C-RAN testbed". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 753–757. DOI: 10.1109/ChinaCom.2012.6417584 (cit. on pp. 17, 41, 42).

[57] J. Li, D. Chen, Y. Wang, and J. Wu. “Performance Evaluation of Cloud-RAN System with Carrier Frequency Offset”. In: *Globecom Workshops (GC Wkshps), 2012 IEEE*. Dec. 2012, pp. 222–226. DOI: 10.1109/GLOCOMW.2012.6477573 (cit. on pp. 18, 42).

[58] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu. “Multi-Cell MIMO Cooperative Networks: A New Look at Interference”. In: *Selected Areas in Communications, IEEE Journal on* 28.9 (2010), pp. 1380–1408. ISSN: 0733-8716. DOI: 10.1109/JSAC.2010.101202 (cit. on pp. 18, 42).

[59] A. Liu and V. Lau. “Joint power and antenna selection optimization for energy-efficient large distributed MIMO networks”. In: *Communication Systems (ICCS), 2012 IEEE International Conference on*. 2012, pp. 230–234. DOI: 10.1109/ICCS.2012.6406144 (cit. on pp. 18, 42).

[60] L. Liu, F. Yang, R. Wang, Z. Shi, A. Stidwell, and D. Gu. “Analysis of handover performance improvement in cloud-RAN architecture”. In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 850–855. DOI: 10.1109/ChinaCom.2012.6417603 (cit. on p. 18).

[61] Wipro Technologies. *Software-Defined Radio Technology Overview, White Paper*. August 2002 (cit. on p. 18).

[62] M. Bansal, J. Mehlman, S. Katti, and P. Levis. “OpenRadio: A Programmable Wireless Dataplane”. In: *ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN’12), ACM SIGCOMM 2012*, pp. 109–114 (cit. on p. 18).

[63] S. Thomas. *Poll: Savings Drive CRAN Deployments*. LightReading, July 2015 (cit. on pp. 19, 41).

[64] *Front-haul Compression for Emerging C-RAN and Small Cell Networks*. Tech. rep. Integrated Device Technology, April 2013 (cit. on p. 21).

[65] W. Huitao and Z. Yong. *C-RAN Bearer Network Solution*. Tech. rep. ZTE, November 2011 (cit. on pp. 21, 27, 41, 42).

[66] S. Namba, T. Warabino, and S. Kaneko. “BBU-RRH Switching Schemes for Centralized RAN”. In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 762–766. DOI: 10.1109/ChinaCom.2012.6417586 (cit. on pp. 22, 42).

[67] H. Raza. “A brief survey of radio access network backhaul evolution: part I”. In: *Communications Magazine, IEEE* 49.6 (2011), pp. 164–171. ISSN: 0163-6804. DOI: 10.1109/MCOM.2011.5784002 (cit. on pp. 23, 25, 42).

[68] C. Chen, J. Huang, W. Jueping, Y. Wu, and G. Li. *Suggestions on Potential Solutions to C-RAN*. Tech. rep. NGMN Alliance, 2013 (cit. on pp. 24, 38, 40, 42, 44, 88).

[69] J. Segel and M. Weldon. *lightRadio Portfolio: White Paper 1*. Tech. rep. Alcatel-Lucent Bell Labs, 2011 (cit. on p. 25).

[70] *3GPP TR 36.932 Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN V 12.1.0*. Mar. 2013 (cit. on pp. 25, 37).

[71] Z. Ghebretensae, K. Laraqui, S. Dahlfort, J. Chen, Y. Li, J. Hansryd, F. Ponzini, L. Giorgi, S. Stracca, and A. Pratt. "Transmission solutions and architectures for heterogeneous networks built as C-RANs". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 748–752. DOI: 10.1109/ChinaCom.2012.6417583 (cit. on pp. 25, 27, 41, 42).

[72] *E-BLINK. Wireless Fronthaul Technology.* [cited: January 2016]. URL: http://e-blink.com/ (cit. on p. 25).

[73] J. H. Lee, S.-H. Cho, K. H. Doo, S.-I. Myong, J. H. Lee, and S. S. Lee. "CPRI transceiver for mobile front-haul based on wavelength division multiplexing". In: *ICT Convergence (ICTC), 2012 International Conference on*. 2012, pp. 581–582. DOI: 10.1109/ICTC.2012.6387205 (cit. on pp. 26, 42).

[74] F. Ponzini, L. Giorgi, A. Bianchi, and R. Sabella. "Centralized radio access networks over wavelength-division multiplexing: a plug-and-play implementation". In: *Communications Magazine, IEEE* 51.9 (2013). ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6588656 (cit. on pp. 26, 42).

[75] *ITU-T G.709/Y.1331, Interfaces for the optical transport network*. Geneva, Feb. 2012 (cit. on p. 27).

[76] *ODU0 and ODUflex A Future-Proof Solution for OTN Client Mapping*. Tech. rep. TPACK, February 2010 (cit. on p. 27).

[77] B. Liu, X. Xin, L. Zhang, and J. Yu. "109.92-Gb/s WDM-OFDMA Uni-PON with dynamic resource allocation and variable rate access". In: *OPTICS EXPRESS, Optical Society of America* 20.10 (May 2012) (cit. on pp. 27, 42).

[78] J. Fabrega, M. Svaluto Moreolo, M. Chochol, and G. Junyent. "WDM overlay of distributed base stations in deployed passive optical networks using coherent optical OFDM transceivers". In: *Transparent Optical Networks (ICTON), 2012 14th International Conference on*. 2012, pp. 1–4. DOI: 10.1109/ICTON.2012.6253934 (cit. on pp. 27, 42).

[79] S. Chia, M. Gasparroni, and P. Brick. "The Next Challenge for Cellular Networks: Backhaul". In: *IEEE Microwave* (August 2009) (cit. on p. 27).

[80] R. Sánchez, L. Raptis, and K. Vaxevanakis. "Ethernet as a Carrier Grade Technology: Developments and Innovations". In: *IEEE Communications Magazine* (September 2008) (cit. on p. 28).

[81] *3GPP TS 36.104 Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception V 12.0.0*. July 2013 (cit. on pp. 28, 30, 84, 87, 124, 125).

[82] *3GPP TS 36.133 Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for Support of Radio Resource Management. V 12.0.0*. July 2013 (cit. on pp. 28, 87).

[83] *NFV Cloud RAN with Ethernet Fronthaul.* [cited: January 2016]. URL: http://www.altiostar.com/solution/ (cit. on pp. 28, 36).

[84] W. Liu, K. Chen, W. Ma, and S. Norberg. "Remote radio data transmission over ethernet". United States Patent Application 20120113972 A1. 2012 (cit. on p. 28).

[85] Altera. *SoftSilicon OTN Processors*. [cited: September 2013]. URL: http://www.altera.com/end-markets/wireline/applications/otn/softsilicon-processors/proc-index.html (cit. on pp. 29, 42).

[86] A. Checko, G. Kardaras, C. Lanzani, D. Temple, C. Mathiasen, L. A. Pedersen, and B. Klaps. *OTN Transport of Baseband Radio Serial Protocols in C-RAN Architecture for Mobile Network Applications*. Tech. rep. MTI Mobile and Altera, March 2014 (cit. on pp. 29, 42, 121, 122).

[87] D. Holberg, Hughes Aircraft Company. "An adaptive digital automatic gain control for MTI radar Systems". 3781882. 1973 (cit. on p. 30).

[88] M. Grieger, S. Boob, and G. Fettweis. "Large scale field trial results on frequency domain compression for uplink joint detection". In: *Globecom Workshops (GC Wkshps), 2012 IEEE*. 2012, pp. 1128–1133. DOI: 10.1109/GLOCOMW.2012.6477737 (cit. on pp. 30, 42).

[89] D. Samardzija, J. Pastalan, M. MacDonald, S. Walker, and R. Valenzuela. "Compressed Transport of Baseband Signals in Radio Access Networks". In: *Wireless Communications, IEEE Transactions on* 11.9 (September 2012), pp. 3216–3225. ISSN: 1536-1276. DOI: 10.1109/TWC.2012.062012.111359 (cit. on pp. 30–32, 42).

[90] B. Guo, W. Cao, A. Tao, and D. Samardzija. "CPRI compression transport for LTE and LTE-A signal in C-RAN". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 843–849. DOI: 10.1109/ChinaCom.2012.6417602 (cit. on pp. 30, 32, 42).

[91] I. D. Technology. *Compression IP for Wireless Infrastructure Applications. Product brief*. [July 2013] (cit. on pp. 30, 32, 42).

[92] J. Lorca and L. Cucala. "Lossless compression technique for the fronthaul of LTE/LTE-advanced Cloud-RAN architectures". In: *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a*. June 2013, pp. 1–9 (cit. on pp. 31, 32, 42).

[93] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz). "Robust and Efficient Distributed Compression for Cloud Radio Access Networks". In: *Vehicular Technology, IEEE Transactions on* 62.2 (February 2013), pp. 692–703. ISSN: 0018-9545. DOI: 10.1109/TVT.2012.2226945 (cit. on pp. 31, 32, 42).

[94] Philippe Sehier et al. *Liaisons, Contributions to 3GPP ETSI on Collaborative Radio/MIMO, ORI Interface, ect.* Tech. rep. NGMN Alliance, 2013 (cit. on p. 33).

[95] Ericsson. *World's first microwave connection between LTE main and remote radio units*. [February 2012]. URL: http://www.ericsson.com/news/1588074 (cit. on pp. 33, 42).

[96] X. Wei, X. Qi, L. Xiao, Z. Shi, and L. Huang. "Software-defined Radio based On Cortex-A9". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 758–761. DOI: 10.1109/ChinaCom.2012.6417585 (cit. on p. 34).

[97] D. Martinez-Nieto, V. Santos, M. McDonnell, K. Reynolds, and P. Carlston. "Digital Signal Processing on Intel® Architecture". In: *Intel® Technology Journal* 13 (1 2009) (cit. on p. 34).

[98] N. Kai, S. Jianxing, C. Kuilin, and K. K. Chai. "TD-LTE eNodeB prototype using general purpose processor". In: *"Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on"*. 2012, pp. 822–827. DOI: 10.1109/ChinaCom.2012.6417598 (cit. on p. 34).

[99] S. Zhang, R. Qian, T. Peng, R. Duan, and K. Chen. "High throughput turbo decoder design for GPP platform". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 817–821. DOI: 10.1109/ChinaCom.2012.6417597 (cit. on p. 35).

[100] Z. Guanghui, N. Kai, H. Lifeng, and H. Jinri. "A method of optimizing the de-Rate Matching and demodulation in LTE based on GPP". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 828–832. DOI: 10.1109/ChinaCom.2012.6417599 (cit. on p. 35).

[101] T. Kaitz and G. Guri. "CPU-MPU partitioning for C-RAN applications". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 767–771. DOI: 10.1109/ChinaCom.2012.6417587 (cit. on p. 35).

[102] H. Duan, D. Huang, Y. Huang, Y. Zhou, and J. Shi. "A time synchronization mechanism based on Software Defined Radio of general-purpose processor". In: *Communications and Networking in China (CHINACOM), 2012 7th International ICST Conference on*. 2012, pp. 772–777. DOI: 10.1109/ChinaCom.2012.6417588 (cit. on p. 35).

[103] N. Omnes, M. Bouillon, G. Fromentoux, and O. Le Grand. “A programmable and virtualized network IT infrastructure for the internet of things: How can NFV SDN help for facing the upcoming challenges”. In: *Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on*. Feb. 2015, pp. 64–69. DOI: 10.1109/ICIN.2015.7073808 (cit. on p. 36).

[104] B. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti. “A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks”. In: *Communications Surveys Tutorials, IEEE* 16.3 (Third 2014), pp. 1617–1634. ISSN: 1553-877X. DOI: 10.1109/SURV.2014.012214.00180 (cit. on p. 36).

[105] *OpenFlow - Open Networking Foundation.* [cited: January 2016]. URL: https://www.opennetworking.org/sdn-resources/openflow (cit. on p. 36).

[106] *The OpenDaylight Platform.* [cited: January 2016]. URL: https://www.opendaylight.org/ (cit. on p. 36).

[107] *Open Platform form NFV (OPNFV). Software.* [cited: January 2016]. URL: https://www.opnfv.org/software (cit. on p. 36).

[108] *OpenStack Open Source cloud Computing Software.* [cited: January 2016]. URL: http://www.openstack.org/ (cit. on p. 36).

[109] *Wind River Titanium Server.* [cited: January 2016]. URL: http://www.windriver.com/products/titanium-server/ (cit. on p. 36).

[110] N. Nikaein, R. Knopp, L. Gauthier, E. Schiller, T. Braun, D. Pichon, C. Bonnet, F. Kaltenberger, and D. Nussbaum. “Demo: Closer to Cloud-RAN: RAN As a Service”. In: *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. MobiCom ’15. Paris, France: ACM, 2015, pp. 193–195. ISBN: 978-1-4503-3619-2. DOI: 10.1145/2789168.2789178. URL: http://doi.acm.org.proxy.findit.dtu.dk/10.1145/2789168.2789178 (cit. on p. 36).

[111] *OpenAirInterface.* [cited: January 2016]. URL: http://www.openairinterface.org/ (cit. on pp. 36, 40).

[112] T. Nakamura, S. Nagata, A. Benjebbour, Y. Kishiyama, T. Hai, S. Xiaodong, Y. Ning, and L. Nan. “Trends in small cell enhancements in LTE advanced”. In: *Communications Magazine, IEEE* 51.2 (2013), pp. 98–105. ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6461192 (cit. on p. 37).

[113] G. Brown. *C-RAN – The Next Generation Mobile Access Platform. LightReading webinar.* 2013 (cit. on p. 39).

[114] Next Generation Mobile Networks. *Project Centralized Processing, Collaborative Radio, Real-Time Computing, Clean RAN System (P-CRAN)*. [cited: February 2013]. URL: http://www.ngmn.org/workprogramme/centralisedran.html (cit. on p. 40).

[115] *Mobile Cloud Networking (MCN) Project*. [cited: April 2013]. URL: https://www.mobile-cloud-networking.eu/ (cit. on p. 40).

[116] *FP7 project High capacity network Architecture with Remote radio heads and Parasitic antenna arrays (HARP)*. [cited: February 2014]. URL: http://www.fp7-harp.eu/ (cit. on p. 40).

[117] *iJOIN. Interworking and JOINt Design of an Open Access and Backhaul Network Architecture for Small Cells based on Cloud Networks*. [cited: Septemebr 2013]. URL: http://www.ict-ijoin.eu/ (cit. on p. 40).

[118] D. Sabella, P. Rost, Y. Sheng, E. Pateromichelakis, U. Salim, P. Guitton-Ouhamou, M. Di Girolamo, and G. Giuliani. "RAN as a service: Challenges of designing a flexible RAN architecture in a cloud-based heterogeneous mobile network". In: *Future Network and Mobile Summit (FutureNetworkSummit), 2013*. July 2013, pp. 1–8 (cit. on p. 40).

[119] *Connectivity management for eneRgy Optimised Wireless Dense networks (CROWD)*. [cited: February 2014]. URL: http://www.ict-crowd.eu/publications.html (cit. on pp. 40, 42).

[120] *Motivation and Vision. iCirrus*. [cited: Septemebr 2013]. URL: http://www.icirrus-5gnet.eu/motivation-and-vision/ (cit. on p. 40).

[121] *ERAN press release. Cloud-based technology for 5G mobile networks based technology for 5G mobile networks*. [cited: January 2016]. URL: http://comcores.com/Files/2014%2012%2018%20ERAN%20press%20release.pdf (cit. on p. 40).

[122] R. Kokku, R. Mahindra, H. Zhang, and S. Rangarajan. "NVS: a virtualization substrate for WiMAX networks". In: *MOBICOM'10*. 2010, pp. 233–244 (cit. on p. 42).

[123] D. Raychaudhuri and M. Gerla. *New Architectures and Disruptive Technologies for the Future Internet: The Wireless, Mobile and Sensor Network Perspective. Technical Report 05-04, GENI Design Document*. 2005 (cit. on p. 42).

[124] L. Xia, S. Kumar, X. Yang, P. Gopalakrishnan, Y. Liu, S. Schoenberg, and X. Guo. "Virtual WiFi: bring virtualization from wired to wireless". In: *VEE'11*. 2011, pp. 181–192 (cit. on p. 42).

[125] G. Smith, A. Chaturvedi, A. Mishra, and S. Banerjee. “Wireless virtualization on commodity 802.11 hardware”. In: *Proceedings of the second ACM international workshop on Wireless network testbeds, experimental evaluation and characterization*. WinTECH ’07. Montreal, Quebec, Canada: ACM, 2007, pp. 75–82. ISBN: 978-1-59593-738-4. DOI: 10.1145/1287767.1287782. URL: http://doi.acm.org/10.1145/1287767.1287782 (cit. on p. 42).

[126] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel. “LTE wireless virtualization and spectrum management”. In: *Wireless and Mobile Networking Conference (WMNC), 2010 Third Joint IFIP*. 2010, pp. 1–6. DOI: 10.1109/WMNC.2010.5678740 (cit. on p. 42).

[127] L. Zhao, M. Li, Y. Zaki, A. Timm-Giel, and C. Gorg. “LTE virtualization: From theoretical gain to practical solution”. In: *Teletraffic Congress (ITC), 2011 23rd International*. 2011, pp. 71–78 (cit. on p. 42).

[128] Z. Zhu, P. Gupta, Q. Wang, S. Kalyanaraman, Y. Lin, H. Franke, and S. Sarangi. “Virtual base station pool: towards a wireless network cloud for radio access networks”. In: *Proceedings of the 8th ACM International Conference on Computing Frontiers*. CF ’11. Ischia, Italy: ACM, 2011, 34:1–34:10. ISBN: 978-1-4503-0698-0. DOI: 10.1145/2016604.2016646. URL: http://doi.acm.org/10.1145/2016604.2016646 (cit. on p. 42).

[129] G. Aljabari and E. Eren. “Virtualization of wireless LAN infrastructures”. In: *Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2011 IEEE 6th International Conference on*. Vol. 2. 2011, pp. 837–841. DOI: 10.1109/IDAACS.2011.6072889 (cit. on p. 42).

[130] H. Coskun, I. Schieferdecker, and Y. Al-Hazmi. “Virtual WLAN: Going beyond Virtual Access Points”. In: *ECEASST* (2009), pp. -1–1 (cit. on p. 42).

[131] Y. Al-Hazmi and H. De Meer. “Virtualization of 802.11 interfaces for Wireless Mesh Networks”. In: *Wireless On-Demand Network Systems and Services (WONS), 2011 Eighth International Conference on*. 2011, pp. 44–51. DOI: 10.1109/WONS.2011.5720199 (cit. on p. 42).

[132] M. Li, L. Zhao, X. Li, X. Li, Y. Zaki, A. Timm-Giel, and C. Gorg. “Investigation of Network Virtualization and Load Balancing Techniques in LTE Networks”. In: *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. 2012, pp. 1–5. DOI: 10.1109/VETECS.2012.6240347 (cit. on p. 42).

[133] G. Bhanage, D. Vete, I. Seskar, and D. Raychaudhuri. “SplitAP: Leveraging Wireless Network Virtualization for Flexible Sharing of WLANs”. In: *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. 2010, pp. 1–6. DOI: 10.1109/GLOCOM.2010.5684328 (cit. on p. 42).

[134] J. Vestin, P. Dely, A. Kassler, N. Bayer, H. Einsiedler, and C. Peylo. "CloudMAC: towards software defined WLANs". In: *ACM SIGMOBILE Mobile Computing and Communications Review* 16.4 (2013), pp. 42–45. ISSN: 15591662, 19311222. DOI: 10.1145/2436196.2436217 (cit. on p. 42).

[135] *Network Functions Virtualisation – Introductory White Paper*. Tech. rep. ETSI, October 2012 (cit. on p. 42).

[136] NFV working group. *Network Function Virtualization – Introductory white paper ETSI*. 2012 (cit. on p. 42).

[137] NFV working group. *Network Function Virtualization; Architectural Framework*. 2013 (cit. on p. 42).

[138] NFV working group. *Network Function Virtualization (NFV); Use Cases*. 2013 (cit. on p. 42).

[139] H. Kim. and G. F. "Improving network management with software defined networking". In: *Communications Magazine, IEEE* 51.2 (2013), pp. 114–119. DOI: 10.1109/MCOM.2013.6461195 (cit. on p. 42).

[140] X. Jin, L. E. Li, L. Vanbever, and J. Rexford. "SoftCell: Scalable and Flexible Cellular Core Network Architecture". In: *Proceedings of the Ninth ACM Conference on Emerging Networking Experiments and Technologies*. CoNEXT '13. Santa Barbara, California, USA: ACM, 2013, pp. 163–174. ISBN: 978-1-4503-2101-3. DOI: 10.1145/2535372.2535377. URL: http://doi.acm.org/10.1145/2535372.2535377 (cit. on p. 42).

[141] A. Gudipati, D. Perry, L. E. Li, and S. Katti. "SoftRAN: software defined radio access network". In: *ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN'13), ACM SIGCOMM 2013*, pp. 25–30 (cit. on p. 42).

[142] K. Pentikousis, Y. Wang, and W. Hu. "Mobileflow: Toward software-defined mobile networks". In: *Communications Magazine, IEEE* 51.7 (2013), pp. 44–53. ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6553677 (cit. on p. 42).

[143] H. Ali-Ahmad, C. Cicconetti, A. de la Olivia, M. Draexler, R. Gupta, V. Mancuso, L. Roullet, and V. Sciancaleporee. "CROWD: An SDN Approach for DenseNets". In: *Second European Workshop on Software Defined Networks (EWSDN), 2013*, pp. 25–31 (cit. on p. 42).

[144] K. Kiyoshima, T. Takiguchi, Y. Kawabe, and Y. Sasaki. "Commercial Development of LTE-Advanced Applying Advanced C-RAN Architecture". In: *NTT DOCOMO Technical Journal* 17.2 (Oct. 2015), pp. 10–18 (cit. on p. 41).

[145] *Common Public Radio Interface (CPRI); Interface Specification V6.1*. July 2014 (cit. on p. 43).

[146] U. Dotsch, M. Doll, H.-P. Mayer, F. Schaich, J. Segel, and P. Sehier. “Quantitative analysis of split base station processing and determination of advantageous architectures for LTE”. In: vol. 18. 1. June 2013, pp. 105–128 (cit. on p. 44).

[147] B. Haberland. *Smart Mobile Cloud*. 2013 (cit. on pp. 44, 51, 52).

[148] *Small cell virtualization functional splits and use cases*. Tech. rep. Small Cell Forum, June 2015 (cit. on pp. 44, 45, 51, 52, 84, 85).

[149] H. Lee, Y. O. Park, and S. S. Song. “A traffic-efficient fronthaul for the cloud-RAN”. In: *Information and Communication Technology Convergence (ICTC), 2014 International Conference on*. Oct. 2014, pp. 675–678. DOI: 10.1109/ICTC.2014.6983252 (cit. on p. 44).

[150] B. Haberland. *Cloud RAN architecture evolution from 4G to 5G Mobile Systems*. 2014 (cit. on p. 44).

[151] E. Dahlman, S. Parkvall, and J. Skold. *4G: LTE/LTE-Advanced for Mobile Broadband*. Elsevier Science, 2011. ISBN: 9780123854902. URL: https://books.google.com.qa/books?id=DLbsq9GD0zMC (cit. on pp. 52, 53).

[152] 3GPP, Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA). *Packet Data Convergence Protocol (PDCP) specification (Release 8)*. Tech. rep. 36.323, v. 8.6.0. 2009 (cit. on p. 51).

[153] 3GPP, Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA). *Radio Link Control (RLC) protocol specification (Release 8)*. Tech. rep. 36.322, v. 8.8.0. 2010 (cit. on p. 51).

[154] 3GPP, Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA). *Medium Access Control (MAC) protocol specification (Release 8)*. Tech. rep. 36.321, v. 8.11.0. 2011 (cit. on p. 52).

[155] C. Desset, B. Debaillie, V. Giannini, A. Fehske, G. Auer, H. Holtkamp, W. Wajda, D. Sabella, F. Richter, M. J. Gonzalez, H. Klessig, I. Gódor, M. Olsson, M. A. Imran, A. Ambrosy, and O. Blume. “Flexible power modeling of LTE base stations”. In: *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. Apr. 2012, pp. 2858–2862. DOI: 10.1109/WCNC.2012.6214289 (cit. on pp. 55–57).

[156] Alcatel-Lucent and Vodafone Chair on Mobile Communication Systems. *Study on Energy Efficient Radio Access Network (EERAN) Technologies*. Unpublished Project Report, Technical University of Dresden, Dresden, Germany, 2009 (cit. on pp. 56, 57).

[157] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. V. D. Perre. “Challenges and enabling technologies for energy aware mobile radio networks”. In: *IEEE Communications Magazine* 48.11 (Nov. 2010), pp. 66–72. ISSN: 0163-6804. DOI: 10.1109/MCOM.2010.5621969 (cit. on p. 56).

[158] Z. Hasan, H. Boostanimehr, and V. K. Bhargava. "Green Cellular Networks: A Survey, Some Research Issues and Challenges". In: *IEEE Communications Surveys Tutorials* 13.4 (Fourth 2011), pp. 524–540. ISSN: 1553-877X. DOI: 10.1109/SURV.2011.092311.00031 (cit. on p. 56).

[159] G. Auer, V. Giannini, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, C. Desset, and O. Blume. "Cellular Energy Efficiency Evaluation Framework". In: *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*. May 2011, pp. 1–6. DOI: 10.1109/VETECS.2011.5956750 (cit. on p. 57).

[160] C. Chen. "The Notion of overbooking and Its Application to IP/MPLS Traffic Engineering". In: *Request for Comments: internet draft <draft-cchen-te-overbooking-01.txt>* (November 2001) (cit. on p. 58).

[161] M. Stasiak, M. Głąbowski, A. Wiśniewski, and P. Zwierzykowski. *Modeling and dimensioning of mobile networks : from GSM to LTE*. John Wiley & Sons Ltd., 2011 (cit. on p. 58).

[162] *OPNET*. [cited: January 2016]. URL: http://www.opnet.com/ (cit. on p. 58).

[163] *Many cities. MIT Senseable City Lab*. [cited: January 2016]. URL: http://www.manycities.org/ (cit. on p. 59).

[164] *Mobility Report*. Tech. rep. Ericsson, November 2013 (cit. on p. 64).

[165] *3GPP TR 25.896 Feasibility Study for Enhanced Uplink for UTRA FDD*. Mar. 2004 (cit. on p. 66).

[166] X. Cheng. "Understanding the Characteristics of Internet Short Video Sharing: YouTube as a Case Study". In: *Procs of the 7th ACM SIGCOMM Conference on Internet Measurement, San Diego (CA, USA), 15*. 2007, p. 28 (cit. on p. 68).

[167] J. J. Lee and M. Gupta. *A new traffic model for current user web browsing behavior*. Tech. rep. Intel, 2007 (cit. on p. 68).

[168] *Average Web Page Breaks 1600K*. [cited: June 2015]. URL: http://www.websiteoptimization.com/speed/tweak/average-web-page/ (cit. on p. 68).

[169] Metro Ethernet Forum. *EVC Ethernet Services Definitions Phase 3, Tech. Spec. MEF 6.2*. July 2014 (cit. on p. 72).

[170] *3GPP TS 36.321 Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification V 12.0.0*. Dec. 2013 (cit. on p. 84).

[171] B. Sadiq, R. Madan, and A. Sampath. “Downlink Scheduling for Multiclass Traffic in LTE”. In: *EURASIP Journal on Wireless Communications and Networking* 2009.1 (2009), p. 510617. ISSN: 1687-1499. DOI: 10.1155/2009/510617. URL: http://jwcn.eurasipjournals.com/content/2009/1/510617 (cit. on p. 84).

[172] T. Pfeiffer and F. Schaich. *Optical Architectures For Mobile Back- And Fronthauling*. OFC/NFOEC wireless backhauling workshop. Los Angeles: Alcatel-Lucent Bell Labs Stuttgart, Mar. 5, 2012 (cit. on p. 84).

[173] J. H. et al. *Further Study on Critical C-RAN Technologies, Version 1.0,* tech. rep. NGMN Alliance, 2015 (cit. on p. 85).

[174] Metro Ethernet Forum. *Mobile Backhaul Implementation Agreement – Phase 2, Amendment 1 – Small Cells, Tech. Spec. MEF 22.1.1*. July 2014 (cit. on pp. 85, 111).

[175] *Federal Communications Commission FCC 15-9, PS Docket No. 07-114*. Feb. 2015 (cit. on p. 86).

[176] M. Weiss. *Telecom Requirements for Time and Frequency Synchronization*. Presented at 52nd Meeting of the Civil GPS Service Interface Committee in Nashville, September 17th-18th, 2012. Time and Frequency Division, NIST (cit. on pp. 86, 87).

[177] E. Metsälä and J. Salmelin. *Mobile Backhaul*. John Wiley and Sons, Ltd., 2012. ISBN: 978-1-119-97420-8 (cit. on p. 86).

[178] *ETSI/GSM. GSM Recommendation 05.10, 3GPP TS 05.10 Radio subsystem synchronization, Release Ph1*. Oct. 1992 (cit. on p. 86).

[179] H. Li. *Consideration on CRAN fronthaul architecture and interface*. Presented at 1st NGFI workshop in Beijing, China, June 4th, 2015. China Mobile Research Institute, June 4, 2015 (cit. on p. 86).

[180] C.-L. I. *NGFI: Next Generation Fronthaul towards 4.5G and 5G*. Presented at 1st NGFI workshop in Beijing, China, June 4th, 2015. China Mobile Research Institute, June 4, 2015 (cit. on pp. 86, 88, 89).

[181] D. Bladsjo, M. Hogan, and S. Ruffini. “Synchronization aspects in LTE small cells”. In: *Communications Magazine, IEEE* 51.9 (2013), pp. 70–77. ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6588653 (cit. on p. 87).

[182] C.-L. I, Y. Yuan, J. Huang, S. Ma, C. Cui, and R. Duan. “Rethink fronthaul for soft RAN”. In: *Communications Magazine, IEEE* 53.9 (Sept. 2015), pp. 82–88. ISSN: 0163-6804. DOI: 10.1109/MCOM.2015.7263350 (cit. on p. 88).

[183] *IEEE 1904.3 Task Force*. URL: http://www.ieee1904.org/3 (cit. on p. 88).

[184] "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems". In: *IEEE Std 1588-2008 (Revision of IEEE Std 1588-2002)* (July 2008), pp. c1–269. DOI: 10.1109/IEEESTD.2008.4579760 (cit. on p. 90).

[185] "IEEE Standard for Ethernet". In: *IEEE Std 802.3-2012* (Dec. 2012), pp. 1–. DOI: 10.1109/IEEESTD.2012.6419735 (cit. on p. 90).

[186] P. A. Smith. *CPRI "FrontHaul" requirements discussion with TSN*. San Diego, CA: Huawei, July 1, 2014 (cit. on p. 91).

[187] *Precise Timing for Base Stations in the Evolution to LTE*. Tech. rep. Vitesse, 2013 (cit. on p. 93).

[188] L. Xie, Y. Wu, and J. Wang. "Efficient time synchronization of 1588v2 technology in packet network". In: *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*. May 2011, pp. 181–185. DOI: 10.1109/ICCSN.2011.6014030 (cit. on p. 95).

[189] "Time Sensitive Networking Task Group of IEEE 802.1, IEEE 802.1Qbv/D2.3 Bridges and Bridged Networks—Amendment: Enhancements for Scheduled Traffic, IEEE". In: (2015) (cit. on p. 100).

[190] "Time Sensitive Networking Task Group of IEEE 802.1, IEEE 802.1Qbu/D2.2 Bridges and Bridged Networks—Amendment: Frame Preemption, IEEE". In: (2015) (cit. on p. 100).

[191] "IEEE Higher Layer LAN Protocols Working Group (C/LM/WG802.1), Time-Sensitive Networking for Fronthaul, P802.1CM PAR". In: (2015) (cit. on p. 102).

[192] "LAN/MAN Standards Committee, IEEE P802.3br/D1.0 Draft Standard for Ethernet Amendment: Specification and Management Parameters for Interspersing Express Traffic., IEEE". In: (2014) (cit. on p. 104).

[193] T. Wan and P. Ashwood. *A Performance Study of CPRI over Ethernet*. Presented at IEEE 1904.3 Task Force meeting in Louisville, January 30th, 2015. Huawei Canada Research Center, Jan. 30, 2015 (cit. on p. 104).

[194] J. Farkas and B. Varga. *Applicability of Qbu and Qbv to Fronthaul*. Ericsson, Nov. 11, 2015 (cit. on p. 105).

[195] T. Stern and K. Bala. *Multiwavelength Optical Networks: A Layered Approach*. Addison-Wesley professional computing series. Addison-Wesley, 1999. ISBN: 9780201309676. URL: https://books.google.dk/books?id=C%5C_lSAAAAMAAJ (cit. on p. 106).

[196] A. De La Oliva, X. Costa Perez, A. Azcorra, A. Di Giglio, F. Cavaliere, D. Tiegelbekkers, J. Lessmann, T. Haustein, A. Mourad, and P. Iovanna. "Xhaul: toward an integrated fronthaul/backhaul architecture in 5G networks". In: *Wireless Communications, IEEE* 22.5 (Oct. 2015), pp. 32–40. ISSN: 1536-1284. DOI: 10.1109/MWC.2015.7306535 (cit. on p. 111).

[197] N. J. Gomes, P. Chanclou, P. Turnbull, A. Magee, and V. Jungnickel. “Fronthaul evolution: From CPRI to Ethernet”. In: *Optical Fiber Technology* 26, Part A (2015). Next Generation Access Networks, pp. 50–58. ISSN: 1068-5200. DOI: http://dx.doi.org/10.1016/j.yofte.2015.07.009. URL: http://www.sciencedirect.com/science/article/pii/S1068520015000942 (cit. on p. 111).