LUT University

MULTI-CONNECTIVITY FOR USER THROUGHPUT ENHANCEMENT IN 5G NON-TERRESTRIAL NETWORKS

Lappeenranta-Lahti University of Technology LUT

Master's Program in Computational Engineering, Master's Thesis

2022

Mikko Majamaa

Examiners: Associate Professor Lassi Roininen
Ph.D. Jani Puttonen

Supervisors: Associate Professor Lassi Roininen
Ph.D. Jani Puttonen

ABSTRACT

Lappeenranta-Lahti University of Technology LUT
School of Engineering Science
Computational Engineering
Technomathematics

Mikko Majamaa

Multi-Connectivity for User Throughput Enhancement in 5G Non-Terrestrial Networks

Master's thesis
2022
42 pages, 27 figures, 2 tables, 1 appendix
Examiners: Associate Professor Lassi Roininen and Ph.D. Jani Puttonen

Keywords: Multi-Connectivity, 5G, Non-Terrestrial Networks, ns-3, Satellite, Communications, SATCOM

Fifth Generation (5G) wireless systems aim to confront with such use-cases as Ultra-Reliable Low Latency Communications (URLLC), massive Machine-Type Communications (mMTC) and enhanced Mobile Broadband (eMBB). To meet the increasing throughput and reliability demands, satellites may be used to complement the 5G Terrestrial Networks (TNs), e.g., in rural areas, in case of emergencies or in areas with peak demands. The EU funded "Dynamic spectrum sharing and bandwidth-efficient techniques for high-throughput MIMO Satellite systems" (DYNASAT) project focuses on research on techniques which aim to improve reliability and throughput in Non-Terrestrial Networks (NTNs). Multi-Connectivity (MC), where a user can be connected to multiple base stations simultaneously, is one of the bandwidth efficient techniques under the research in the project. In this master's thesis, the focus is on MC to improve users' experienced throughput. First, a study of relevant specifications and algorithms is performed. Then, MC feature is implemented to a 5G NTN simulator. The simulation results indicate that there exist scenarios where using MC in 5G NTNs is beneficial. The significance of this thesis lies in the fact that there exist no service providers that offer system level simulations for 5G NTNs with MC support and packet-level precision.

TIIVISTELMÄ

Lappeenrannan-Lahden teknillinen yliopisto LUT
School of Engineering Science
Laskennallinen tekniikka
Teknillinen matematiikka

Mikko Majamaa

Moniyhteydellisyys käyttäjän tiedonsiirtonopeuden parantamiseksi 5G satelliittiverkoissa

Diplomityö
2022
42 sivua, 27 kuvaa, 2 taulukkoa, 1 liite
Tarkastajat: Apulaisprofessori Lassi Roininen ja FT Jani Puttonen

Hakusanat: Moniyhteydellisyys, 5G, Ei-maanpäällinen verkko, ns-3, Satelliitti, Satelliittikommunikointi, SATCOM

Viidennen sukupolven (5G) langattomat järjestelmät pyrkivät vastaamaan sellaisiin tarpeisiin kuin erittäin luotettava, matalan viiveen viestintä, massiivinen koneiden välinen viestintä sekä parannettu mobiililaajakaista. Satelliittiverkkoja voidaan käyttää täydentämään maanpäällisiä verkkoja, jotta kasvavat tiedonsiirto-sekä luotettavuustarpeet saadaan tyydytettyä esim. maaseuduilla, onnettomuuksien sattuessa tai alueilla, joilla on korkea tiedonsiirtotarve. EU-rahoitteinen DYNASAT-projekti keskittyy tutkimaan tekniikoita, joilla pyritään kasvattamaan yhteyksien luotettavuutta ja tiedonsiirtokykyä satelliittiverkoissa. Moniyhteydellisyys, jossa päätelaite voi olla yhteydessä useaan tukiasemaan samanaikaisesti, on yksi projektissa kaistanleveyden tehokkuuuden parantamiseksi tutkittu keino. Tässä diplomityössä tutkimusaiheena on moniyhteydellisyys käyttäjien kokeman tiedonsiirron parantamiseksi. Ensin suoritetaan kirjallisuuskatsaus olennaisiin määrityksiin sekä algoritmeihin. Sen jälkeen kehitetään moniyhteydellisyysominaisuus 5G satelliittiverkkosimulaattoriin. Simulaatiotulokset osoittavat, että on skenaarioita, joissa moniliitännäisyydestä on hyötyjä 5G satelliittiverkoissa. Tämän diplomityön merkittävyys on siinä, että ei ole toimijoita, jotka tarjoaisivat 5G satelliittiverkkosimulaatioita moniyhteydellisyystuella sekä pakettitason tarkkuudella.

ACKNOWLEDGEMENTS

The subject of this thesis was provided by Magister Solutions Ltd. The implementation and the produced reports to the DYNASAT project during my daily work was used in the thesis. I would like to thank the company for this opportunity. I would also like to thank my supervisors, associate professor Lassi Roininen from LUT University and Dr. Jani Puttonen from the company, for their professional guidance and the latter also for his insights regarding the research on MC. Moreover, I am grateful for all the generous help I have received from my co-workers (special thanks to Dr. Henrik Martikainen for our weekly MC planning meetings and other support), for without the help, it would have been impossible to carry out this thesis.

Jyväskylä, March 16, 2022

Mikko Majamaa

LIST OF ABBREVIATIONS

3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
5GC	5G Core
AHP	Analytic Hierarchy Process
AMF	Access and Mobility Management Function
CA	Carrier Aggregation
CC	Component Carrier
CDF	Cumulative Distribution Function
CN	Core Network
DC	Dual Connectivity
DSA	Dynamic Spectrum Allocation
DYNASAT	"Dynamic spectrum sharing and bandwidth-efficient techniques for high-throughput MIMO Satellite systems"
E-UTRA	Evolved Universal Terrestrial Access
E2E	End-to-End
eMBB	enhanced Mobile Broadband
en-gNB	en-Next Generation Node B
eNB	Evolved Node B
EPC	Evolved Packet Core
GEO	Geostationary Orbit
gNB	Next Generation Node B
gNB-CU	gNB-Centralized Unit
gNB-DU	gNB-Distributed Unit
HetNet	Heterogeneous Networks
IoT	Internet of Things
LAN	Local Area Network
LEO	Low Earth Orbit
MAC	Media Access Control
MC	Multi-Connectivity
MCG	Master Cell Group
mMTC	massive Machine-Type Communications
MN	Master Node
MR-DC	Multi Radio-Dual Connectivity

ng-eNB	Next Generation eNB
NG-RAN	Next Generation Radio Access Network
NGSO	Non-Geostationary Orbit
NR	New Radio
ns-3	Network Simulator 3
NTN	Non-Terrestrial Network
PDCP	Packet Data Convergence Protocol
PGW	Packet Network Data Gateway
PoC	Proof-of-Concept
RA	Random Access
RAN	Radio Access Network
RAT	Radio Access Technology
RF	Radio Frequency
RRC	Radio Resource Control
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
SCG	Secondary Cell Group
SGW	Serving Gateway
SINR	Signal-to-Interference-plus-Noise Ratio
SLS	System Level Simulator
SN	Secondary Node
SRS	Sounding Reference Signal
TCP	Transmission Control Protocol
TN	Terrestrial Network
TR	Technical Report
TS	Technical Specification
UDP	User Datagram Protocol
UE	User Equipment
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communications
VR	Virtual Reality

CONTENTS

1 INTRODUCTION

1.1 Background

5G wireless systems aim to tackle with such use-cases as URLLC, mMTC and eMBB [1]. Some concrete examples from these use-cases include, but are not limited to, telesurgery, Internet of Things (IoT) and Virtual Reality (VR). To meet the increasing throughput and reliability demands, satellites may be used to complement the 5G TNs, e.g., in rural areas, in the case of emergencies or in areas with peak demands. Especially Non-Geostationary Orbit (NGSO) satellite systems have been under intense research activities in the recent years because of the low latency associated with them, advancements in technology and relatively cheap price. These systems are being deployed by such companies as Amazon, Telesat and SpaceX, among many others. In Figure 1, typical interworking of a TN-NTN system is presented. In the depicted scenario, the satellites involved are Low Earth Orbit (LEO) and Geostationary Orbit (GEO) satellites. LEO satellites can provide services that are delay sensitive, whereas GEO satellites can provide additional bandwidth.

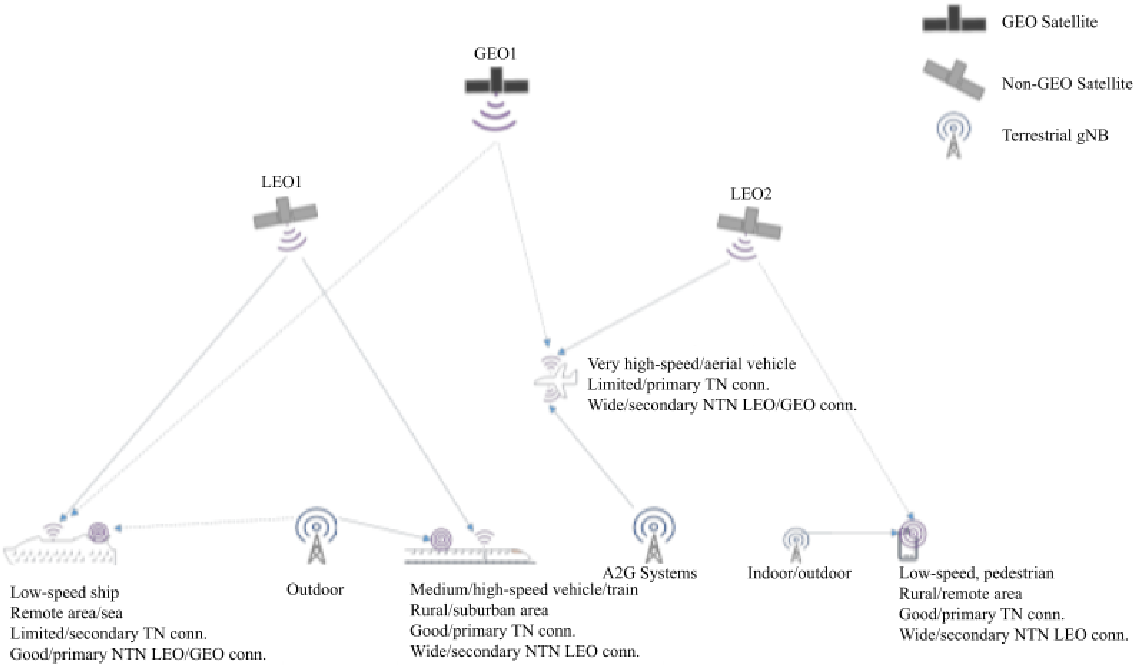


Figure 1. Typical interworking of a TN-NTN system. [2]

Moreover, European Commission has ambitious goals towards future communications requirements [3]. A satellite component may play a key-role to meet these requirements.

The EU funded DYNASAT [4] project focuses on research on spectrum sharing and bandwidth efficient techniques in NTNs. These techniques aim to improve reliability, as well as throughput. MC, where a user can be connected to multiple base stations simultaneously, is one of the bandwidth efficient techniques under the research in the project.

Magister Solutions Ltd's focus on the project is to study the spectrum sharing and bandwidth efficient techniques, namely Dynamic Spectrum Allocation (DSA) and MC, with the means of system simulations. In this master's thesis, the focus is on the MC part, specifically to improve users' experienced throughput. First a study of relevant specifications and algorithms is performed. Then, MC feature is implemented to the company's 5G NTN simulator. The significance of this thesis lies in the fact that there exist no operators that provide system level simulations for 5G NTNs [5] with MC support and packet-level precision.

1.2 Objectives and delimitations

The objective of this master's thesis is to conduct a review on MC related architectures and algorithms in specifications and research literature. After the review, the objective is to implement MC feature to the company's 5G NTN simulator. The work in this thesis also includes testing and implementing two different Secondary Node (SN) addition algorithms, but does not include testing or implementing different traffic steering algorithms.

1.3 Structure of the thesis

The structure of this report is as follows. In Chapter 2, relevant architectures to MC in specifications are presented. In Chapter 3, MC related algorithms are reviewed. Description of the implementation of the MC feature to the simulator is given in Chapter 4. Following the implementation description, in Chapter 5, simulation scenarios and results are presented. In Chapter 6, the work that has been done, and that is planned to do in the future, is discussed. Finally, the work is concluded in Chapter 7.

2 MULTI-CONNECTIVITY AND RELATED ARCHITECTURES IN SPECIFICATIONS

3rd Generation Partnership Project (3GPP) [6] is a standardization organization which aims to provide specifications for mobile communications. Among other contributions, it publishes Technical Specifications (TSs) and Technical Reports (TRs), which correspond to specifications and results of some preliminary study on some subject, respectively. The organization was founded in 1998 first to provide specifications for Third Generation (3G) telecommunications. Later it has released specifications for all of the following generations. History of different telecommunications generations can be found in Figure 2.

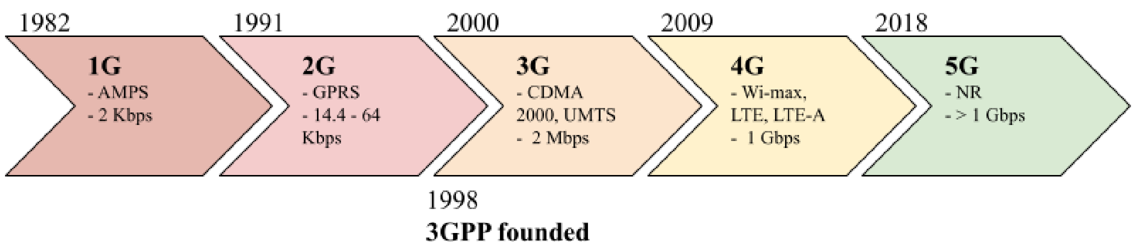


Figure 2. Different telecommunications generations from 1G to 5G and example technologies associated with them, as well as bandwidths they may offer [7], [8], [9].

In 3GPP Release 17 [10], to be published in March 2022, there's a work item called "Solutions for NR to support non-terrestrial networks (NTN)" [2]. This work item aims to study the needed (if any) enhancements for the 5G TN specifications to be applicable for NTNs. For example, MC is specified for TNs in TS 37.340 (Release 15) [11] but is not yet specified for NTNs. Since the lack of specifications and research results available related to MC with NTNs, much of the following sections use specifications and research from TNs as a reference. Furthermore, one of the planned key contributions of the DYNASAT project is to use its results in relevant specification activities.

Next, the most relevant parts of the 3GPP specifications/reports related to this work are reviewed. These consist of technical details of MC, radio protocol architectures, network interfaces and transparent/regenerative payload satellites. The details described here are essential for the implementation of the MC feature to the 5G NTN System Level Simulator (SLS).

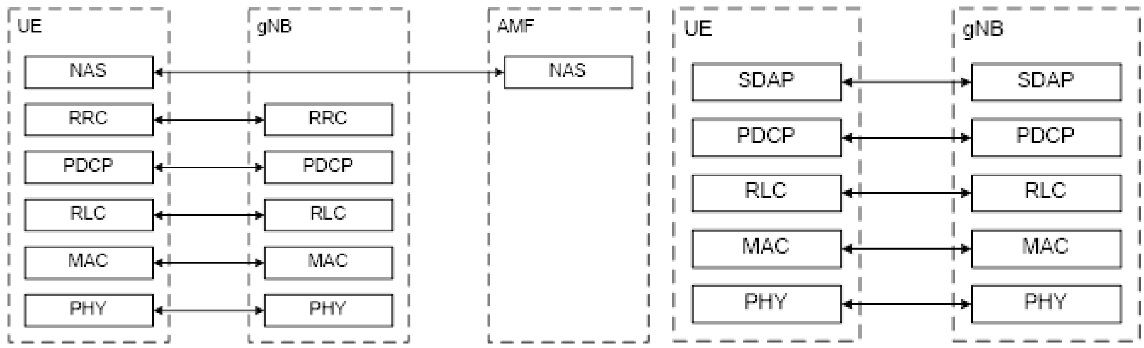
2.1 Generalities

MC is a technique where a User Equipment (UE), e.g., a mobile phone, can be connected to multiple base stations simultaneously. It can be used to enhance throughput and/or reliability e.g., UE being at a cell edge where throughput does not fulfil the requirements or UE having a demand for URLLC. URLLC consists of, but is not limited to, telesurgery and automated traffic. In [12], MC scheduling is divided into three categories: load balancing, packet duplication and packet splitting (also a form of load balancing). In load balancing, the channel is selected out of the multiple channels that the UE is connected to in such a such way that the data traffic between the channels is evenly balanced. In packet duplication packets are duplicated and sent over all the channels. In packet splitting, the packet data is split, and different parts of the packet may be sent over different channels. One can come to conclusion that packet duplication would be beneficial especially in URLLC communications and load balancing in throughput enhancement, though, both can serve to enhance both types of communication requirements.

MC is a generalization of Dual Connectivity (DC), where a UE can be connected to two radio base stations simultaneously, i.e., Master Node (MN) and SN. In MC, there could be multiple SNs. The reader should note that in some parts of the following sections, it is talked about DC as it is discussed as such in the reference specifications, papers, and reports. In the DYNASAT project, the focus of interest is in MC, so when confronting the term DC, the reader may extend to thinking of multiple SNs. Furthermore, in the project, asynchronous MC is considered, i.e., in the data transmission, the sub-frame boundaries are not synchronized between the MN and SN(s) [13], because of possibly too high synchronization delays between the NTN nodes.

2.2 Definitions

To understand the different types of DCs that are found in the specifications, it is important to first explain some important terms that are to be used. Evolved Universal Terrestrial Access (E-UTRA) corresponds to Fourth Generation (4G) wireless communication link technology, whereas New Radio (NR) is the corresponding technology in 5G. Evolved Packet Core (EPC) is the Core Network (CN) entity associated with 4G that provides connectivity between the different parts of the network, e.g., Local Area Networks (LANs) or subnetworks. The corresponding CN in the case of 5G is 5G Core (5GC). Moreover, a protocol stack defines operations related to communications. Control plane is responsible for control messages and user plane for user data traffic. Each layer consists of different operations, e.g., responsibilities of Radio Resource Control (RRC) layer include addition, modification and release of SNs and connection related signalling in general, whereas physical (PHY) layer is responsible for actual transmission of data. Media Access Control (MAC) layer offers data transfer and radio resource allocation services to upper layers of the protocol stack. Protocol stack of control plane and user plane for 4G/5G wireless communications between a user and a base station is depicted in Figure 3. Even though the stacks coincide with 4G and 5G, the operations at the different layers may differ.



(a) (b)

Figure 3. Protocol stack of: (a) Control plane (b) User plane for 4G/5G wireless communications between a user (UE) and a base station (gNB). [14]

Multi Radio-Dual Connectivity (MR-DC), as specified in TS 37.340 [11], is a generalization of E-UTRA intra DC, as described in TS 36.300 [15]. In E-UTRA intra DC, a UE can be connected to two radio base stations that are Evolved Node Bs (eNBs) operating in EPC. Whereas, MR-DC is supported with EPC, as well as with 5GC. In MR-DC, a UE can be connected to an MN and an SN, where one node provides NR access and the other either NR or E-UTRA access. An example MR-DC architecture is depicted in Figure 4, where the nodes operate in EPC, and nodes are providing both NR and E-UTRA accesses. These nodes are called eNBs, i.e., radio base stations that provide E-UTRA user and control plane terminations towards the UE [15] and en-Next Generation Node Bs (en-gNBs), i.e., a radio base stations that provide NR user and control plane terminations towards the UE. In the case of EPC (as in Figure 4) eNB acts as an MN and en-gNB as an SN.

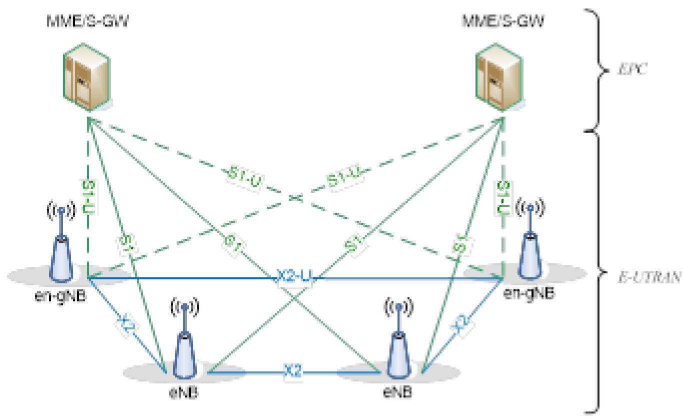


Figure 4. Example MR-DC architecture. [11]

In the case of MR-DC with 5GC, there are two different radio base stations operating: Next Generation eNBs (ng-eNBs), i.e., a radio base stations providing E-UTRA user and control plane terminations for the UE and Next Generation Node Bs (gNBs) providing NR user and control plane terminations towards the UE [14].

Carrier Aggregation (CA) [16] is a technique to increase bandwidth where two or more Component Carriers (CCs) are aggregated. CA can also be used in parallel with MR-DC. For example, an NR node (i.e., en-gNB or gNB), can aggregate at maximum 16 CCs for downlink and 16 CCs for uplink [11]. Moreover, in [17], the authors explain that using CA is based on the same Radio Resource Management (RRM) measurement design as in MC: in CA, additional CCs are set up based on measurement reports from the UE. Rather than at the Packet Data Convergence Protocol (PDCP) layer (as in the MC case), in CA the carriers can be aggregated at the MAC layer. In general, MC could be thought of as distributed CA where the aggregated CCs are aggregated from different nodes.

2.3 Scenarios and architectures described by 3GPP

2.3.1 Radio protocol architecture

Relevant radio protocol architectures are specified in [11]. In MR-DC, the UE has a single RRC state (the RRC state managed by the MN), as well as a single control plane connection towards the CN. Control plane architecture in the case of 5GC for NR-DC is presented in Figure 5. In this figure, 5GC case is considered. In many of the following figures in this section the approach is similar. What changes in case of EPC and other types of DCs (e.g., EN-DC, NR-EN-DC) is mostly the interfaces between the nodes.

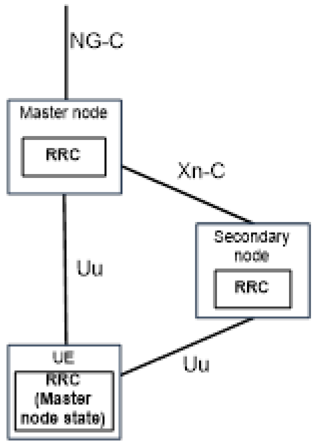


Figure 5. Control plane architecture in the case of 5GC for NR-DC. [11]

From the UE perspective, there are three different bearers in MR-DC: Master Cell Group (MCG), Secondary Cell Group (SCG) and split bearer. In MCG and SCG bearers, only MN or SN radio resources are involved, respectively. In split bearer, radio resources of both nodes are involved. In Figure 6, different bearer options from UE perspective with 5GC is presented. In case of EPC the options are similar, excluding that the MCG PDCP can be either NR or E-UTRA PDCP.

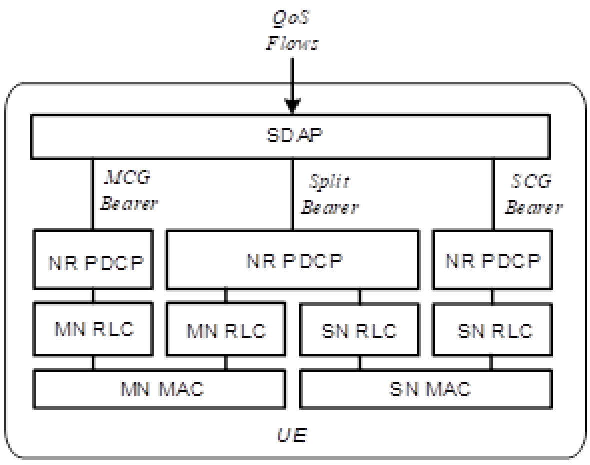
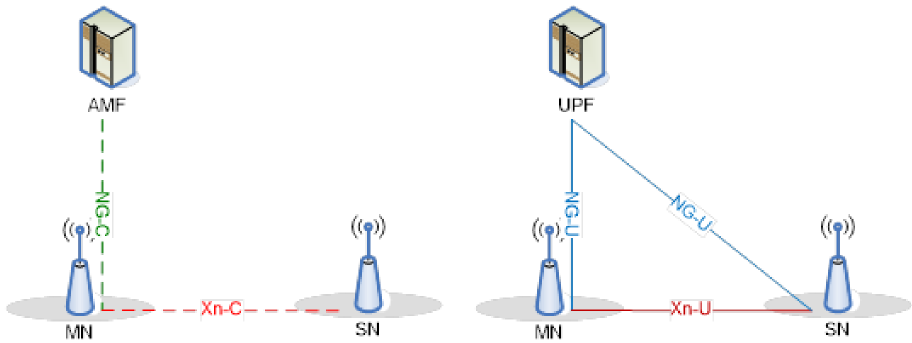


Figure 6. Different bearer options from the UE perspective with 5GC. [11]

2.3.2 Network interfaces

In MR-DC, the MN is connected to the CN entity. The MN and SN are connected through an interface for control signaling and coordination. The MN and SN are primarily responsible for their radio resource allocation. Control and user plane connectivity are presented in Figure 7.



(a) (b)

Figure 7. Connectivity for NR-DC with 5GC in the case of: (a) Control plane (b) User plane. [11]

2.3.3 Transparent and regenerative payloads

As described in TR 38.821 [2], in case of transparent payload satellites, the satellite repeats the signal, corresponding to an analogue Radio Frequency (RF) repeater. Whereas, with regenerative payloads, (part of) base station (e.g., gNB) capabilities are on-board of a satellite, e.g., demodulation, decoding, re-modulation and re-coding functionalities. Next Generation Radio Access Network (NG-RAN) architecture of transparent and regenerative payload satellites are presented in Figures 8 and 9, respectively.

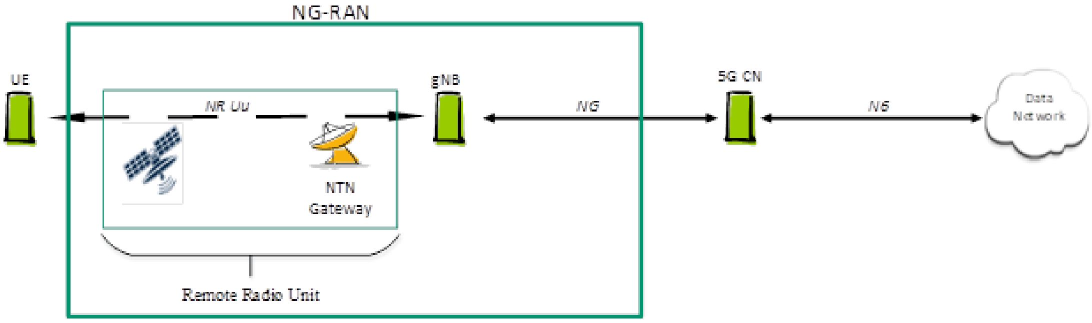


Figure 8. Transparent payload satellite. [2]

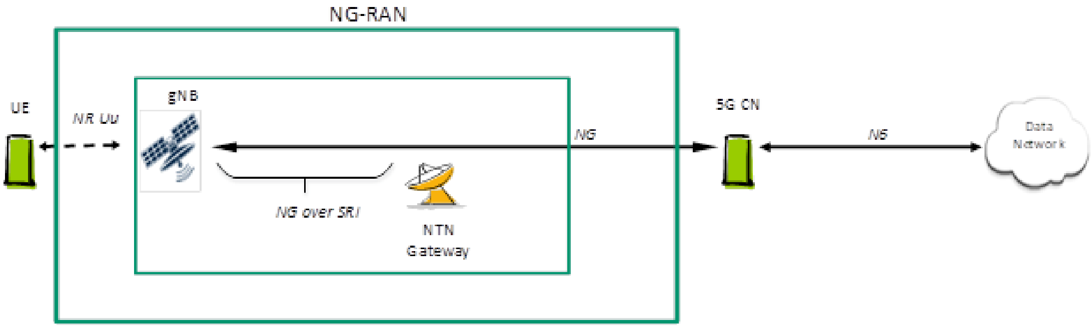


Figure 9. Regenerative payload satellite. [2]

Furthermore, in the same report, a logical split for NG-RAN in the case of regenerative satellites is considered, where a gNB RAN functionalities can be divided to gNB-Centralized Unit (gNB-CU) and gNB-Distributed Units (gNB-DUs). This split is further detailed in TS 38.401 [18] and is depicted in Figure 10. For example, in [19], suitable split option for user plane is common PDCP and RRC for control plane, i.e., these layers in the protocol stack are hosted by the gNB-CU and lower layers by the gNB-DU(s).

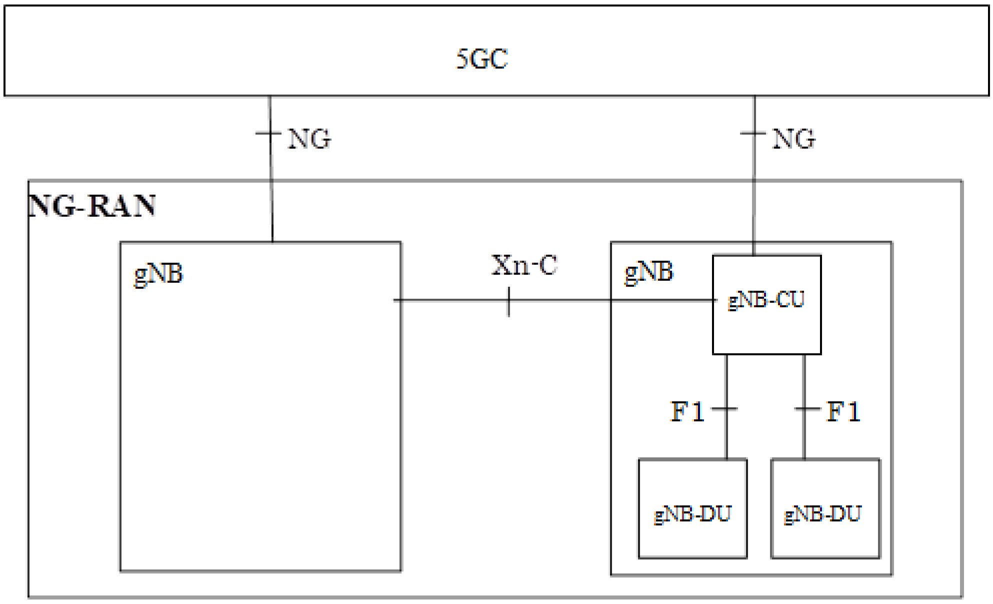


Figure 10. gNB-CU/gNB-DU split. [18]

2.3.4 TN-NTN/NTN-NTN scenarios

TR 38.821 [2], discusses some scenarios, where MC could be beneficial in the case of TN-NTN and NTN-NTN scenarios, i.e., MC in the case of a terrestrial and a non-terrestrial node and in the case of two non-terrestrial nodes. An example scenario for TN-NTN case would be adding an NTN node to underserved areas, e.g., to cell edges, to improve user throughput and/or reliability. More scenarios are presented in TS 22.261 [20]. Example use cases for NTN-NTN scenarios are not provided but can be deduced: think of an area where no TN is available, and a UE is in the edge of two beams of (a) satellite(s).

Radio Access Network (RAN) is a part of the network that implements Radio Access Technology (RAT), e.g., E-UTRA or NR, and provides user an access through a node, e.g., a satellite or a cellurar base station, to the CN. In the following subsections, interworking of TN-NTN/NTN-NTN with different RAN scenarios considering transparent/regenerative payload satellites and cellular base stations are provided. In this thesis, closer interest is on transparent payload LEO satellites because of the project requirements.

2.3.4.1 TN-NTN/NTN-NTN scenarios with transparent payload satellites

The first case, as presented in Figure 11, to consider involves a transparent payload satellite and a cellular base station.

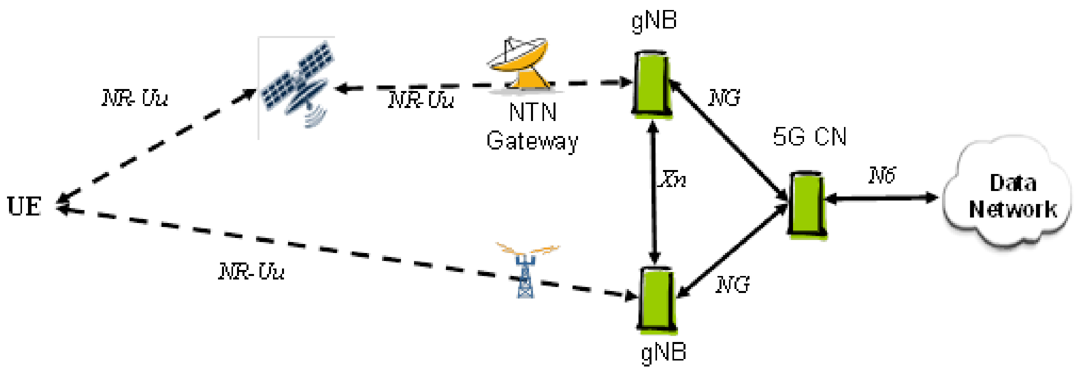


Figure 11. MC involving a transparent payload satellite and a cellurar base station. [2]

The next case considers MC involving two transparent payload satellites (see Figure 12).

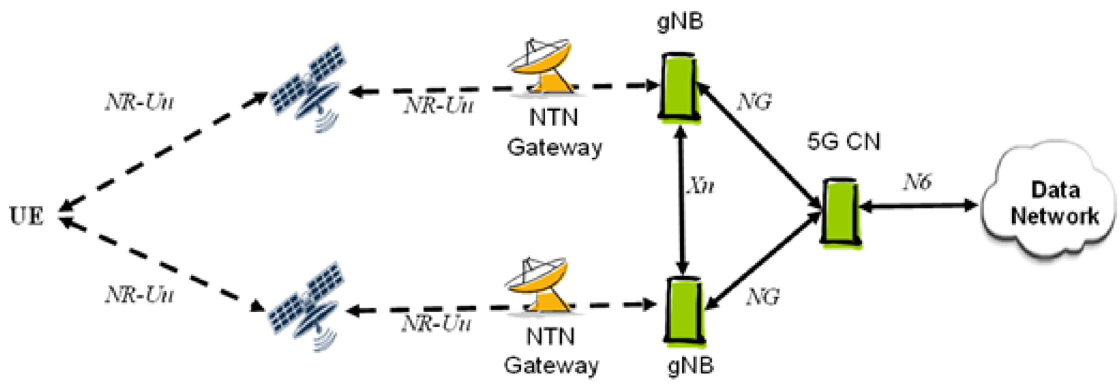


Figure 12. MC involving two transparent payload satellites. [2]

2.3.4.2 TN-NTN/NTN-NTN scenarios with regenerative payload satellites

In Figure 13, MC involving a regenerative satellite with gNB-DU on-board and a cellular base station is depicted.

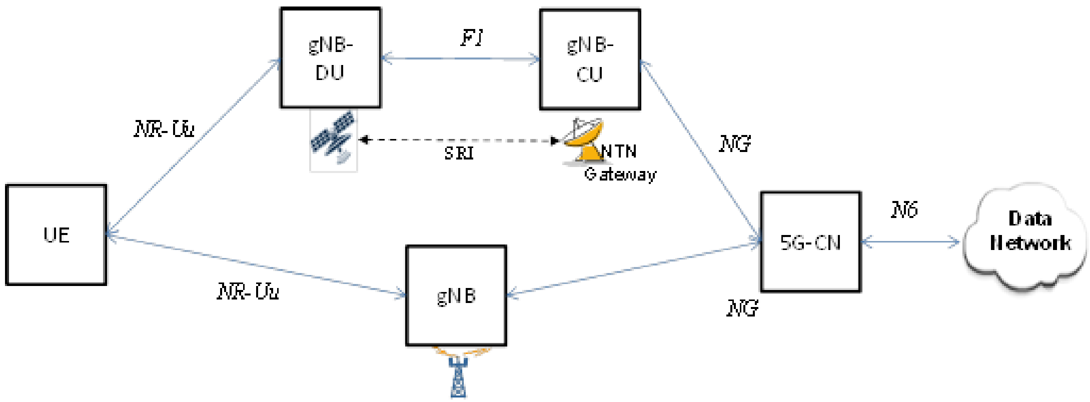


Figure 13. MC involving regenerative payload satellite with gNB-DU on-board and a cellular base station. [2]

Note that, MC between regenerative satellite with gNB on-board and a cellular base station is not addressed because the transport of Xn protocol over the Feeder link (based on Satellite Radio interface) is for further study.

Figure 14 addresses MC involving two regenerative satellites with gNB on-board.

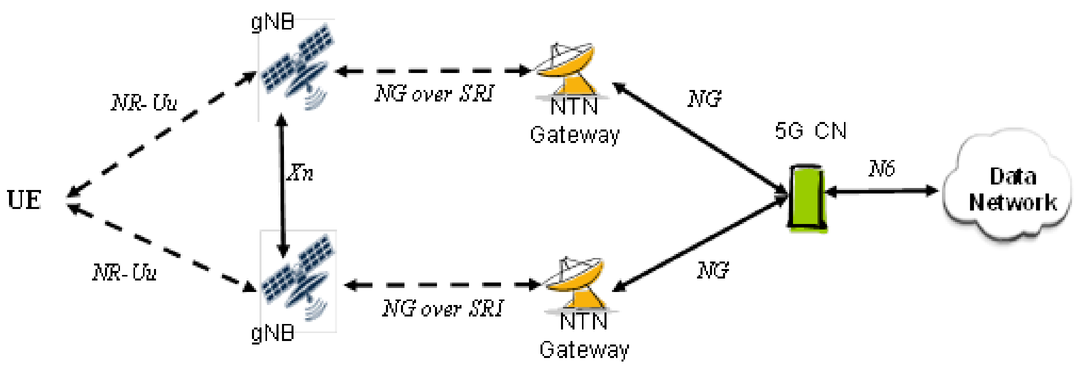


Figure 14. MC involving two regenerative satellites with gNB on-board. [2]

Note also that MC may also involve two regenerative payload satellites with gNB-DUs on-board but is not depicted here.

3 REVIEW OF MULTI-CONNECTIVITY RELATED ALGORITHMS

Algorithmic considerations for MC include i) whether to enable MC for a UE and ii) how to split the data traffic between the MN and the SN. The first point is called cell association. In single-connectivity it is simply choosing the optimal connection for the UE, but in MC it also includes selection of SN(s). The second point is also called traffic steering. Both are wide research areas. In what follows, several cell association and traffic steering algorithms are introduced.

3.1 Cell association algorithms

Authors in [21] consider DC in Heterogeneous Networks (HetNet) with 5G, where a UE can be connected to a macro and a small cell. Enablement of DC for a UE, that has a small cell as an MN, is done by formula:

$$\mathrm{RSRP_s} < \mathrm{RSRP_m} - \mathrm{CRE} + \mathrm{DCrange},$$

where $RSRP_s$ and $RSRP_m$ are the Reference Signal Received Powers (RSRPs) from the small and macro cell, respectively, CRE is the cell range extension (parameter to favour single-connectivity to a small cell) and $DCrange$ is the logical DC area. This enablement of DC is depicted in Figure 15. This sort of configuration corresponds to what is done in [22], also. The same authors as in [22], present a more novel method for configuration of DC in [23], which is a modified opportunistic cell association algorithm. In [24], the authors consider dynamic MC activation for URLLC. As a base, it uses a threshold value to enable MC, but also takes into consideration that MC does not always need to be activated even though the threshold would be fulfilled. The algorithm stores latency budget for the users to keep track of the urgency to activate MC. MC for URLLC is considered, but the algorithm might be modifiable for throughput enhancement also. Authors in [25] propose a cell association algorithm considering MC where each user maintains a list of possible cells in a preferable order by distance and tries to subscribe to a cell in that order. The cell then either accepts the subscription if the cell is not over occupied or if the subscription would lead to a better estimated throughput than with the worst current subscription to the cell, discarding the worst subscription. Five different algorithms for cell association are considered in [26], namely max bitrate, energy efficient max bitrate, max Signal-to-Interference-plus-Noise Ratio (SINR), Analytic Hierarchy Process (AHP) and max clustered bit rate. The algorithms are tested through simulations. AHP, which takes into account performance, robustness and power consumption, performs the best in terms of throughput enhancement with around 75% improvement while max bitrate and energy efficient max bitrate improve around 70% of throughput.

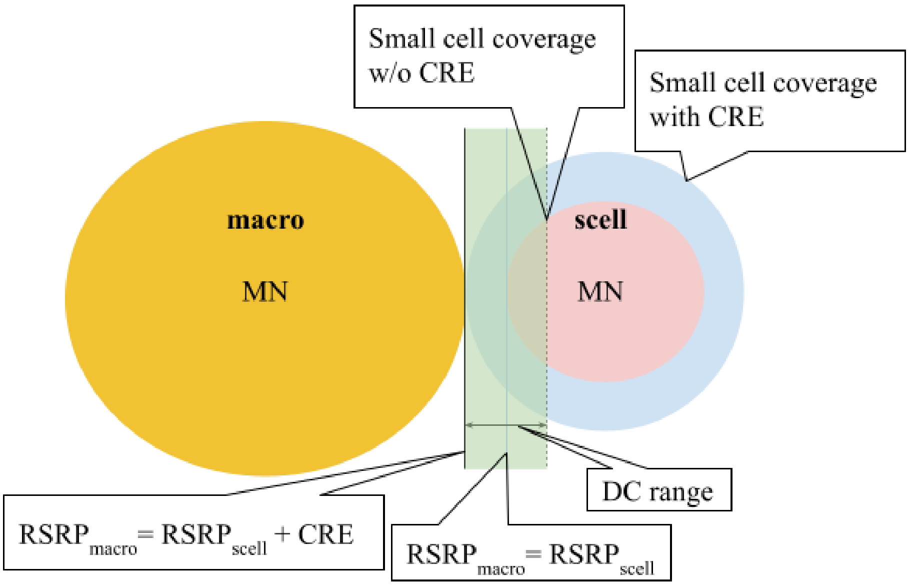


Figure 15. Configuration of DC. Adapted from [21].

Authors in [27] consider DC between 4G and 5G communications. This is useful since high frequencies of 5G can cause the links to be susceptible to failures, e.g. when confronting obstacles. Moreover, this kind of DC may be considered in the deployment phase of 5G when the base stations providing 5G are still sparsely available. In the paper, a simulation framework [28], extended to support the DC feature, for Network Simulator 3 (ns-3), is detailed. In the implementation, each UE is primarily connected to a 4G base station, i.e., eNB. Then, by Sounding Reference Signal (SRS) measurements in the uplink direction, the best available 5G base station, i.e., gNB, is chosen as an SN. The data traffic is always directed through the gNB (if available). Their implementation of DC maintains two links towards the UE, but the traffic is directed only through one of them. Maintaining two links towards a UE allows switching a link faster to another e.g., in the case of a link failure, avoiding use of hard handover where the connection might be lost completely for a while. Even though the type of DC is not exactly similar of what is the interest in this thesis, still the implementation is particularly interesting since it relies partly on the same frameworks and the simulator as the implementation in this thesis, as will be shown later.

3.2 Traffic steering algorithms

Data split between the MN and SN is studied in [22]. Data is forwarded from MN to SN per-request basis, where the SN requests data from the MN that is to be sent to the UE. The amount of the data to be requested is based on pending data requests, scheduled throughput for the UE in the SN and the buffer status of the SN. The traffic control algorithm they developed is also used by authors in [29] where MC in cloud and distributed HetNet architectures is evaluated through simulations. Multiple load balancing algorithms and their mathematical formulation was considered in 5G-ALLSTAR project [30], [31], where interworking of a TN-NTN system is considered. Algorithms include using reinforcement learning, Wardrop Equilibrium based control, Friend or Foe Q-Learning, Games Theory, Linear Programming and AHP. Finally, in [32], the Wardrop Equilibrium control based algorithm, that is chosen in 5G-ALLSTAR for final Proof-of-Concept (PoC) [33], is further described. The wireless network simulator, that is used to evaluate the algorithm, is provided in [34]. For the use-case of this thesis (and the project), the simulator is too simplified since it models mostly downlink and uplink allocations, but not the actual data transmission or protocols in a detailed enough manner.

Authors in [35], formulate the data split problem into a binary integer problem, in the case of a 5G NR TN network. The problem is then modified to be computationally feasible by first maximizing the number of served users and only after that the resources are divided between them. In [36], the problem is also formulated into a binary integer problem, in the case of millimeter-wave networks. The problem is attacked by partitioning the problem into a master problem and a pricing problem. First, in the partition problem the number of possible connection configurations are reduced and then the master problem is solved. LaSR [37] is an MC scheduler whose main advantage, according to the authors, is the ability to take into account realistic system constraints, e.g., discrete and heterogenous ranges of modulations and resource blocks, delays when switching on/off RATs and channel unavailabilities. With Lagrange duality it is possible to solve the original non-convex optimization problem as a set of smaller optimization problems. Benefits are said to be adaptiveness and the lack of need of stochastic traffic assumptions. MC at MAC layer is considered because of finer granularity. Loads of the SNs and RSRPs are considered in the traffic steering problem solution in [38]. The SNs inform the MN about their load statuses as well as their RSRPs after the SN addition has been completed. Then out of these values, by scoring each connection, the MN computes the partition that it directs to each of the SNs.

3.3 Summary

Even though there are multiple more advanced options for secondary cell association, using RSRP measurements with a combination of a threshold value seems to be a popular option. This is likely due to its simplicity without the requirement of the knowledge of the global state of the system and good-enough performance. Still, there exists multiple heuristic/greedy approaches that seem to perform well in secondary cell addition. The traffic steering algorithms that rely on the global knowledge of the system might be problematic because of the high propagation delays. Though, in the case of transparent payload satellites, where the gNB functionalities are anyways on the ground, the problem is mitigated. Secondary cell addition could be considered a somewhat simpler problem, at least when it is considered completely separately from the traffic steering problem, because of the lower number of unknowns. However, many traffic steering algorithms also consider, somewhat implicitly, secondary cell addition.

4 IMPLEMENTATION

4.1 5G NTN Simulator

Magister Solutions Ltd has developed a 5G NTN extension [5] to ns-3 [39]. The extension can be used in system level simulations of standardization processes (e.g., in 3GPP Release 17 [40], as well as in later releases). It can also be used e.g., in testing of radio resource management algorithms, different parameterizations and scenarios. Users may utilize it through Magister Simlab [41], [42]. In the DYNASAT project, the 5G NTN simulator is used to simulate the spectrum sharing and bandwidth efficient transmission techniques under study, e.g., DSA and MC. The latter being the target of interest in this thesis.

ns-3 is an open-source (downloadable from the project website [39]) discrete event network simulator, written in C++, for internet systems. The simulator is mostly used for educational and research purposes. Some common use-cases include studying new network topologies, different parameterizations of networks and new protocols. It can be used to simulate networks with technologies such as Wi-Fi, WiMax, LTE and 5G, just to mention a few. Users may add new modules to the simulator. 5G LENA [43] is one of such modules, which is evolvement of the work done before in LENA [44], the LTE/EPC Network Simulator. According to [5], 5G LENA simulator was chosen as the starting point of the 5G NTN Simulator development, because of the availability, maintenance, community support, and previous experience and competence. The NTN extension includes key features to model non-terrestrial networks. Some of these include modelling of antennas and movement of satellites. Furthermore, different user environments (e.g., urban and suburban), propagation delay model and frequency-division MAC scheduler are also implemented, again, just to mention a few important. Figure 16 lays out the high-level components that form the 5G NTN SLS.

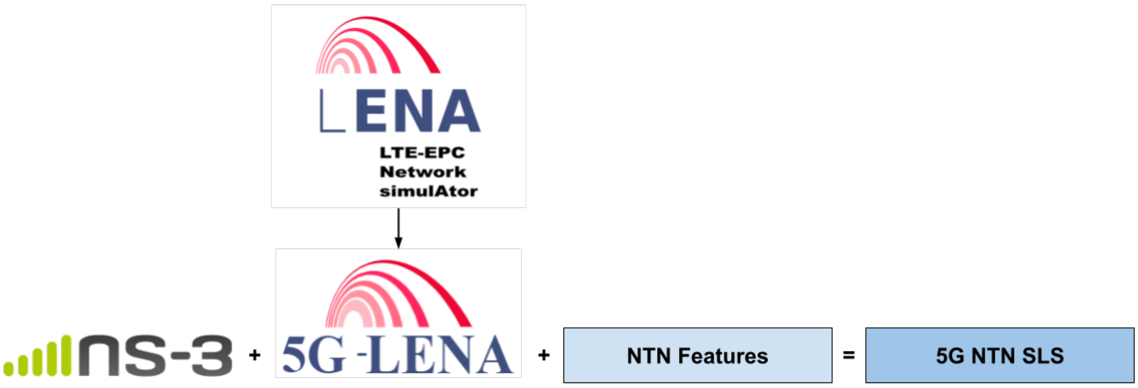


Figure 16. High-level components of the 5G NTN simulator.

In the first phase of implementing MC, most of the changes needed to the architecture relate to the 5G LENA simulator parts of the code. Its End-to-End (E2E) architecture can be found in Figure 17, where the dark grey blocks depict parts of the code that are unmodified ns-3 and LENA components, whereas the light grey parts are the 5G NR features. In the 5G NTN simulator, gNB as seen in Figure 17, is modelled to as a satellite's beam. A transparent payload satellite can then be modelled with propagation delay that considers the user and the feeder links.

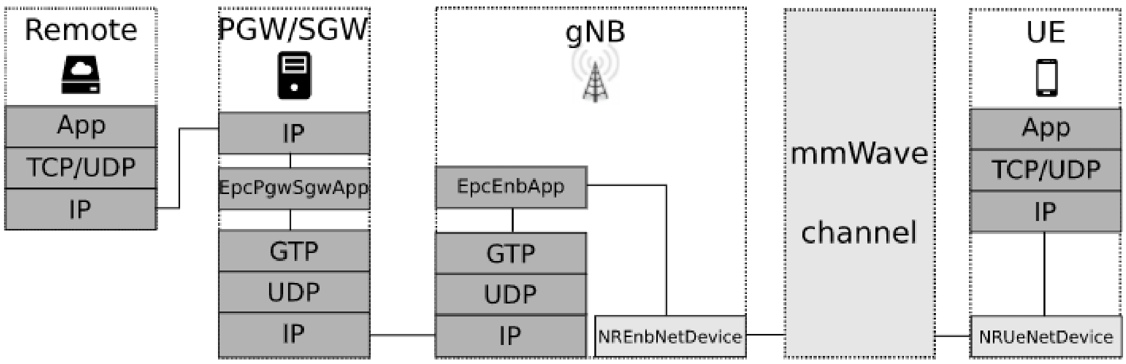


Figure 17. 5G LENA E2E architecture. [43]

The data traverses in the network as follows. For example, from a remote host to a UE first from the remote host to Packet Network Data Gateway (PGW)/Serving Gateway (SGW), then to gNB and finally through wireless channel to UE. From a UE to a remote host the data traverses in the opposite direction. Furthermore, in each of these entities the data passes through the entity's protocol stack, where the data traverses through the stack, is operated at each layer, and finally transmitted to the next entity. The operations at different layers include e.g., routing, scheduling, fragmentation and re-ordering. In Figure 18, more detailed net device ("NREnbNetDevice/NrUeNetDevice" in Figure 17) architecture is depicted. In practice, the figure can be interpreted as follows: there is a satellite (gNB on the left) that transmits data to a user (UE on the right). Between them resides a wireless data channel. Each of the devices have a protocol stacks for the data transmissions. Each layer in the stack has a different function (as explained earlier in the report).

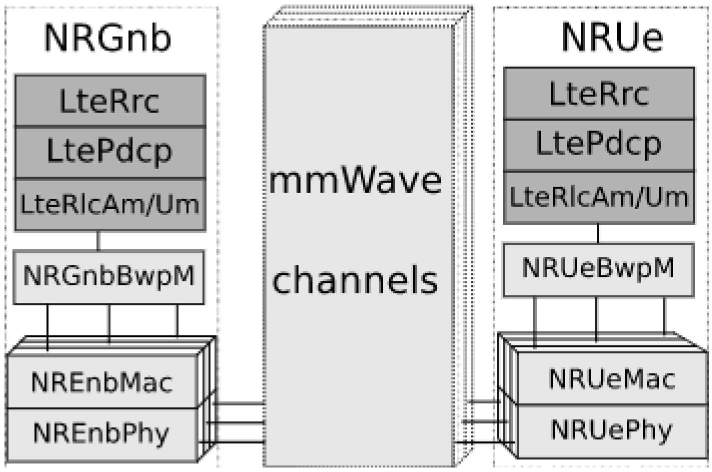


Figure 18. 5G LENA net device architecture. [43]

4.2 Multi-Connectivity extension

In the simulator, MC is implemented at the RRC and PDCP layers, for control plane and user plane, also known as data plane, respectively. Both, the MN and the SN(s), have their own RRC state for each UE that is connected to them, whereas the UE only has a single RRC state (that of the MN's), as described in TS 37.340 [11].

The SN addition process, as specified in TS 37.340 [11], is depicted in Figure 19. The process is initiated by the MN: it sends an SN Addition Request message to the desired SN, which in response replies with an acknowledgement. MN then sends an RRC reconfiguration message to the UE, which does the needed reconfigurations to be able to start receiving data from the SN. The UE responds after the reconfiguration is completed to the MN which then indicates that to the SN. This is the process when no errors occur. Also, note that MC is not yet specified for NTNs by 3GPP so specifications for TRs are used as a reference.

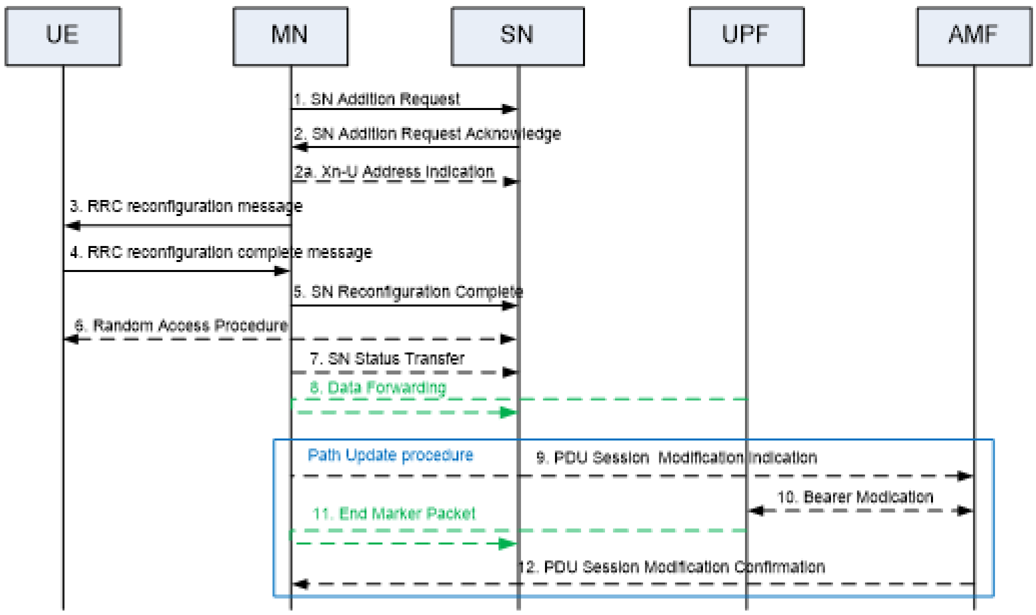


Figure 19. The SN addition process. [11]

The CN functionalities, namely Access and Mobility Management Function (AMF) and User Plane Function (UPF), as seen in the figure are not of interest, at least at this point, in the study of MC and are not modelled in the simulator. Furthermore, Random Access (RA) procedure in MC, i.e., UE forming connection to the SN through RA, is omitted for now for simplicity. In the simulator RA only delays forming of the connection and thus has no significant effect in the research of MC.

Code wise, most of the SN addition procedure takes place in the RRC parts of the LTE module. Figure 20 presents the updated architecture of 5G LENA with MC. Changes to the original E2E architecture include SN that is connected to the MN through X2 interface and to the UE through wireless channel. In the first iteration of the implementation, DC is considered for simplicity, but the modification to the SLS to add multiple SNs is straightforward. Here it is assumed that the MN and the SN operate at different frequency bands which implies that their transmissions do not interfere with each other. It also implies that different protocol stacks at the UE side are required to receive data from different gNBs, which is the most fundamental change required to the net device architecture. In the figure, these stacks refer to the MCG and SCG bearers, as described in TS 37.340 [11].

To the SLS, SN addition algorithms based on Event A3 and A4 [45] have been implemented. Event A3 corresponds to case where a neighbour cell becomes an offset better than the serving cell, whereas Event A4 corresponds to case where the neighbour cell

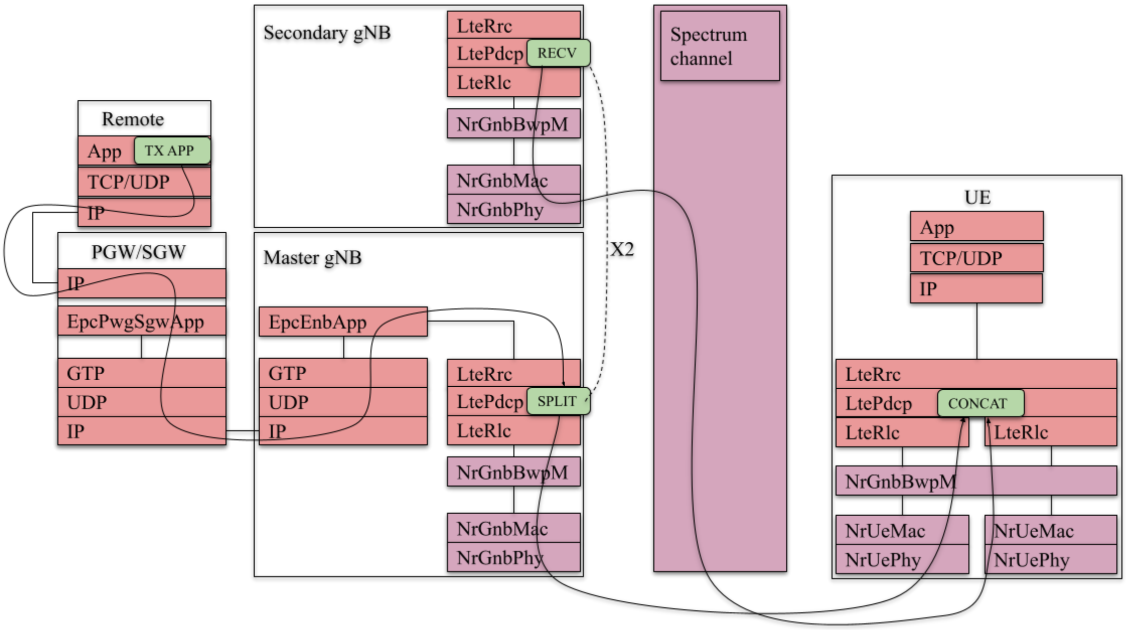


Figure 20. 5G LENA architecture with MC. The red blocks are the ns-3 and LENA components, whereas the purple parts are the 5G NR features. The arrow shows the data traversal in the downlink direction.

becomes better than a threshold. Justification for the implementation of these algorithms is that they seem to be widely used, RSRP measurements are already implemented in the SLS, and they are simple for the first iteration on getting results on possible benefits on using MC. As in the architecture without MC, the data that is to be sent to the UE is first directed to the MN's (the gNB that the UE is attached to in the single-connectivity case) net device. MN can then make the decision, whether to send the data to the UE or redirect it at the PDCP layer to the SN through the X2 interface which then sends the data to the UE. In the current implementation, the split is done evenly, i.e., every other packet is sent by the MN and every other is directed to the SN and sent by it to the UE. More novel traffic steering algorithms are out of the scope of this thesis and are left for future R&D activities. In this work, MC for downlink is considered.

5 SIMULATIONS

The simulation scenario consists of one satellite with two non-overlapping beams operating at different frequencies. Rationale behind the selection of this scenario are twofold. First, it can be used to clearly show that the throughput can be improved when the unutilized second beam is used as an SN for the UEs. Second, when testing more advanced MC algorithms that bring their own complexities to the simulations and their results' analysis, a simple scenario can reduce unnecessary complexity. In the scenarios' traffic model, a remote host sends User Datagram Protocol (UDP) traffic, that traverses through the satellite, to the UEs. 2000 UDP packets, each of size 100 B is sent to each user within 2 s of simulation time (after 1 s of warmup time). The amount of traffic is chosen in such a way that overloading of a beam can be controlled by the amount of UEs. The simulations are run in five drops, i.e., with different random number generator seeds used in the generation of random values in the simulations, e.g., where the UEs are placed. Running the simulations in different drops can help to investigate average behaviour. If instead the simulations were run in a single drop, the results could be misleading.

TR 38.821 [2] provides calibration cases for SLSs. In the simulations, case 10 from Table 6.1.1.1-9 that uses satellite parameter set 1 (see Table 6.1.1.1-1), is considered. This corresponds to a case where the UE is a handheld device, and the satellite is in the low earth orbit. Some important simulation parameters are listed in Table 1.

Table 1. Some important simulation parameters.

Parameter	**Value**
Satellite orbit	LEO 600 km
Beam bandwidth [MHz]	5
Traffic demand per UE [kbps]	800
UE Antenna Type	Handheld

Next, simulations are run when MC is turned off and on with SN addition algorithms based on Event A3 and A4. Two cases are considered 1) heterogeneous load and 2) homogeneous load.

5.1 Heterogeneous Load

In the first case, 20 stationary UEs are randomly placed within 30 km of the center of the satellite's first beam. This is to get the other beam more loaded than the other. It is assumed that it is known that the other beam is overloaded and thus cannot be added as an SN. Initial connection to a cell is done based on UE measurements. Figure 21 depicts the scenario, for one of the drops, after the initial cell selection has been completed.

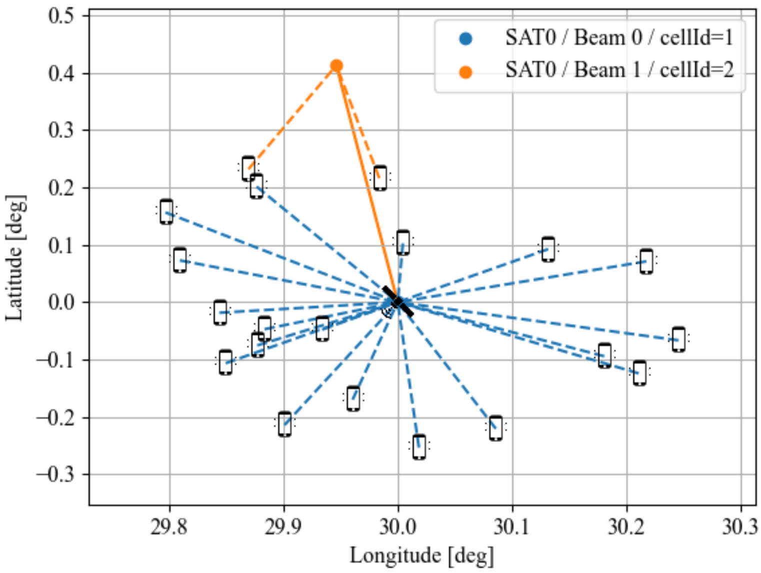


Figure 21. Simulation scenario with a single satellite with two beams where one of the beams is overloaded.

Simulations are first run to find out the RSRP values related to each cell after the initial cell selection. From these values the Cumulative Distribution Functions (CDFs) of offset values (used as a parameter in Event A3-based SN addition algorithm), as well as the RSRP of the secondary cell (used as a parameter in Event A4-based SN addition algorithm), are inspected to determine what the values should be to trigger SN addition for a certain proportion of UEs. Figure 22 illustrates how the use of different offset values with Event A3-based SN addition algorithm effects users' throughputs. From the lowest offset value to the highest these correspond to triggering SN addition on for roughly from 60% down to 10% of UEs with ten percentage point jumps. From the graph it can be observed that using an offset value of -9.16 dBm gives the best throughput enhancement and is thus used in the final simulations related to this scenario.

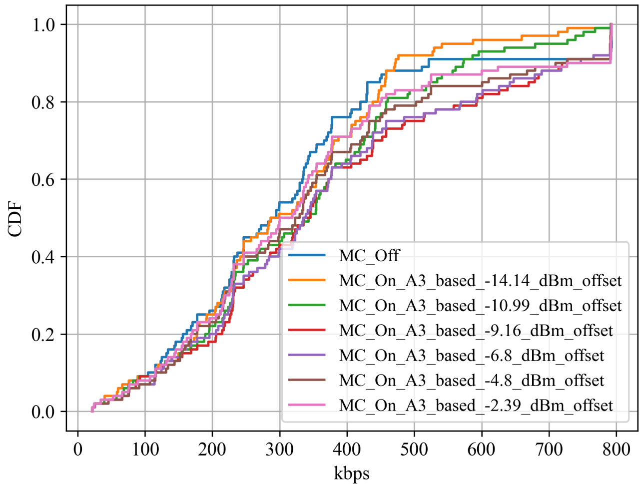


Figure 22. Effect on users' throughputs of using different offset values with Event A3-based SN addition algorithm when one of the beams is overloaded.

The RSRP threshold, -124.48 dBm, that is used with the Event A4-based SN addition algorithm is determined similarly. The results of using SN addition algorithms based on Event A3 and A4, as well as when MC is turned off, is depicted in Figure 23. It shows the CDF of the users' throughputs for all of the drops. From the figure it can be observed that most of the users experienced enhancement in their throughput and that the total throughput of the system increased, as expected.

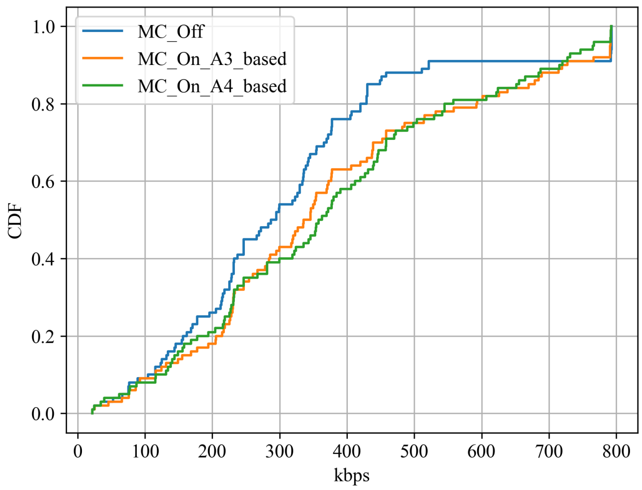


Figure 23. Users' throughputs in the scenario with a single satellite with two beams where one of the beams is overloaded. SN addition algorithms based on Event A3 and A4 compared with the case where MC is turned off.

Figure 24 shows the average throughputs of the simulation drops for each of the users. The results shown in the figure can be found in numerical format from Appendix 1, Table A1.1. With both of the SN addition algorithms, the node with id 14 experienced highest throughput improvement with 98% and 116% in case of Event A3 and A4 based SN addition algorithms used, respectively. Node with id 9 had the least benefit when MC was turned off: its throughput decreased 40% and 33% when using Event A3 and A4 based SN addition algorithms, respectively. This can be the case for example when both of the connections are bad and it is still tried to send through both of them which may lead to dropping a lot of packets. On average, the throughput for a user improved 21.5% and 22.9% with the different algorithms used. The total throughput of the system was 6286.9 kbps when MC was turned on. When MC was turned on and Event A3 based SN addition algorithm was used, the throughput was 7500.1 kbps. When Event A4 based SN addition algorithm was used, it was 7549.6 kbps. The variance in the throughput between the UEs can be explained by their different locations in the area of the beams and the differing Modulation and Coding Schemes (MCS) UEs use depending on the signal strength. MCS defines how many bits is sent in a resource element. Note that the system is overloaded such that not all the data is received by the UE during the simulation, neither with MC turned off nor on. Relatively low throughputs can be explained by looking at the simulation parameters (see Table 1): with handheld devices, experienced SINRs are low. Addition to that, low beam bandwidth limit the experienced throughputs even more.

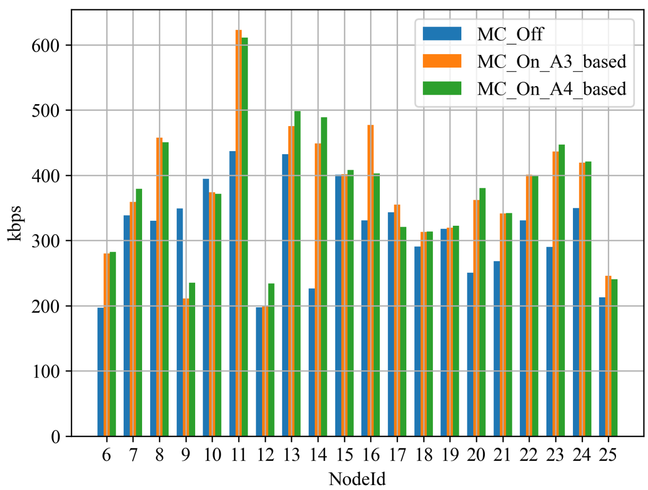


Figure 24. Downlink application call throughput (kbps) per UE, 2000 UDP packets of 100 B sent to each UE.

Another KPI of interest is the percentage of the capacity demand actually served. In the simulations, the total actual capacity demand was 32000 kb (2000 packets × 100 B × 20 users). When MC was turned off, the actual capacity served was 12573.8 kb (2 s × 6286.9 kbps), whereas when MC was turned on, it was 15000.2 kb (2 s × 7500.1 kbps) with Event A3 based SN addition algorithm and 15099.2 kb (2 s × 7549.6 kbps) with Event A4 based SN addition algorithm. So, the percentage of the capacity demand actually served went up from 39.3% to 46.9% and 47.2% when using Event A3 and A4 based SN addition algorithms, respectively.

5.2 Homogeneous Load

In this case, ten stationary UEs are randomly placed within 25 km of the center of each of the beams. Figure 25 depicts the scenario.

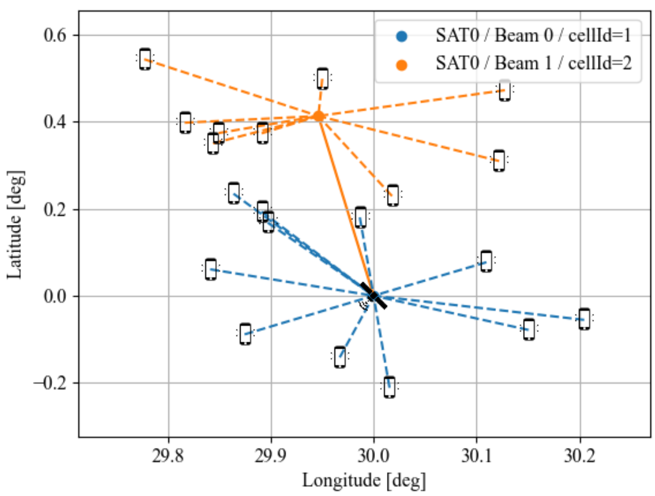


Figure 25. Simulation scenario with a single satellite with two non-overloaded beams with homogeneous load.

The parameters to use with the SN addition algorithms are determined as in the previous scenario. Figure 26 shows, in case of the Event A3-based SN addition algorithm, the effect of different offset values to the throughput. The offsets used correspond to adding an SN from 60% down to 10% of UEs. Choosing a lower offset value deteriorates the throughput since none of the cells are overloaded and adding SNs only generates excess traffic, e.g., in the X2 interfaces.

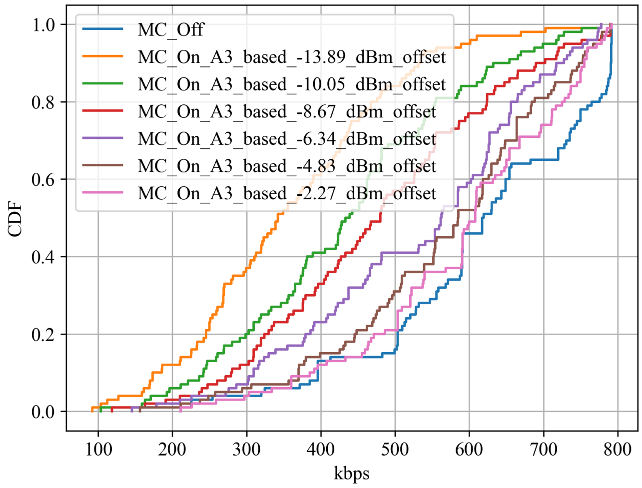


Figure 26. Effect on users' throughputs of using different offset value with the Event A3-based SN addition algorithm when the load is homogeneous.

The best values, by this kind of inspection, for the threshold and offset values are determined in this scenario too. Offset value to be used is -2.27 dBm and threshold -115.85 dBm for the SN addition algorithms based on Event A3 and A4, respectively. This corresponds to turning on MC for roughly 10% of the UEs, i.e., the minimum amount of the cases that are inspected. The results can be found in Figure 27. It can be deduced that using MC in this kind of scenario may even be counterproductive.

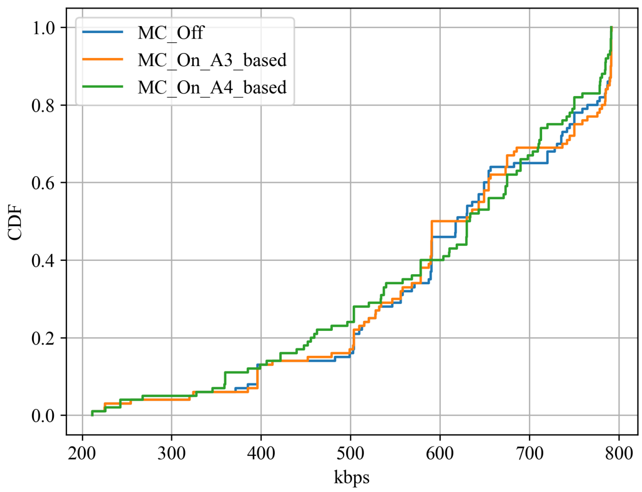


Figure 27. Users' throughputs in the scenario with a single satellite with two non-overloaded beams with homogeneous load. SN addition algorithms based on Event A3 and A4 compared with the case where MC is turned off.

5.3 Summary

Two scenarios with heterogeneous and homogeneous load on a satellite's two beams were inspected through simulations. In the heterogeneous case, where one of the beams was overloaded, users' experienced throughputs enhanced by more than 20% with both of the SN addition algorithms used. As expected, when the beams are not overloaded, using MC can deteriorate performance of the system. Two SN addition algorithms were tested, namely ones that are based on Event A3 and A4. These two scenarios show that more information is needed for better performance than mere RSRP measurements, i.e., load indication whether the cell can handle more UEs. Moreover, the Event A3-based SN addition algorithm itself might be insufficient: if the primary connection is bad, adding an SN that has a signal strength close to equal as bad might be pointless.

6 DISCUSSION

As can be observed from the section regarding the implementation of MC, MC feature has been implemented to the simulator. Given this fact and the literature review done, the research questions set for this thesis were fulfilled. In the previous section, two scenarios with heterogeneous and homogeneous loads were inspected. It was shown that with the given parameters and the heterogeneous scenario, throughput of the users can be enhanced by turning on MC. Furthermore, it was also shown that in the particular homogeneous case that was studied, no enhancement in users' throughput was experienced, as expected. Moreover, two different SN addition algorithms were implemented and tested, namely Event A3 and A4 based. Both of these provided enhancement in users' throughput in the heterogeneous case. The Event A3-based SN addition algorithm itself might be insufficient: if the primary connection is bad, adding an SN that has a signal strength close to equal as bad might be pointless. The Event A4 based SN addition algorithm is a candidate for further studies. The flow between the primary and the master connection was divided evenly. Future R&D activities include research on more novel flow control mechanisms. Also, the need for SN addition for a specific user must be taken into consideration when evaluating SN addition, as well as the capacity available of the SN to be added.

Some other important features to be implemented include support for MC in uplink, definition of more diverse simulation scenarios (e.g., larger satellite constellations and TN/NTN scenarios) and support for more complex traffic models, e.g., ones operating on top of Transmission Control Protocol (TCP). Addition of multiple SNs could also be considered. Several other features might be implemented that are not of utmost importance at this point. These features include e.g., RA process to form the connection between the UE and the SN, MC operation related aspects (e.g., SN release and modification, for more, see TS 37.340 [11]) and including possibility for errors in the SN addition signaling process.

The author's personal contribution regarding the work described in this report include the literature review and the implementation of the MC feature to the 5G NTN SLS (with great support in personal/team discussions). The parametrizations and definitions of the simulation scenarios were conducted together with Dr. Henrik Martikainen. The original version of Figure 20 depicting the architecture of the simulator with MC was made by Dr. Jani Puttonen and was updated by the author according to the final implementation.

7 CONCLUSION

In this thesis, a literature review of MC related algorithms and architectures was conducted. The review was followed by an implementation part where MC feature was added to a 5G NTN simulator. Two SN addition algorithms were implemented, namely Event A3 and A4 based. The simulation results indicate that there exist scenarios where using MC in 5G NTNs is beneficial. The feature can be used in future research activities in algorithm development and testing, as well as standardization work. Furthermore, it can be utilized by the company's simulations as a service product. The significance of this thesis lies in the fact that there exist no service providers that offer system level simulations for 5G NTNs with MC support and packet-level precision.

REFERENCES

[1] 5G Standalone Architecture. Technical White Paper, Samsung, 2021.

[2] TR 38.821: Solutions for NR to support Non-Terrestrial Networks (NTN). V16.0.0, January 2020.

[3] Connectivity for a Competitive Digital Single Market - Towards a European Gigabit Society. European Commission, Directorate-General for Communications Networks, Content and Technology, 2021. Accessed on: Dec. 17, 2021. [Online]. Available: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52016DC0587.

[4] About DYNASAT. 2021. Accessed on: Dec. 17, 2021. [Online]. Available: https://www.dynasat.eu/about-dynasat/.

[5] Jani Puttonen, Lauri Sormunen, Henrik Martikainen, Sami Rantanen, and Janne Kurjenniemi. A System Simulator for 5G Non-Terrestrial Network Evaluations. *2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 292–297, 2021.

[6] 3GPP. Accessed on: Dec. 17, 2021. [Online]. Available: https://www.3gpp.org/.

[7] Xavier Jover Segura. Heterogeneous Wireless Networks QoE Framework. Doctoral dissertation, University College London (UCL), 2015.

[8] Jyoti Goyal, Khushboo Singla, Akashdeep, and Sarbjeet Singh. A Survey of Wireless Communication Technologies from 1G to 5G. *International Conference on Computer Networks and Inventive Communication Technologies*, pages 613–624, 2020.

[9] Opeoluwa Tosin Eluwole, Nsima Udoh, Mike Ojo, Chibuzo Okoro, and Akintayo Johnson Akinyoade. From 1G to 5G, what next? *IAENG International Journal of Computer Science*, 45:413–434, 2018.

[10] 3GPP Release 17. Accessed on: Dec. 17, 2021. [Online]. Available: https://www.3gpp.org/release-17.

[11] TS 37.340: NR; Multi-connectivity; Overall description; Stage-2. V16.7.0, September 2021.

[12] Marie-Theres Suer, Christoph Thein, Hugues Tchouankem, and Lars Wolf. Multi-Connectivity as an Enabler for Reliable Low Latency Communications—An Overview. *IEEE Communications Surveys Tutorials*, 22(1):156–169, 2020.

[13] Kunihiko Teshima Toru Uchino and Kazuki Takeda. Carrier Aggregation Enhancement and Dual Connectivity Promising Higher Throughput and Capacity. *NNT Docomo Technical Journal*, 17, 2015.

[14] TS 38.300: NR; NR and NG-RAN Overall description; Stage-2. V16.7.0, September 2021.

[15] TS 36.300: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2. V16.6.0, June 2021.

[16] Carrier Aggregation explained. Accessed on: Nov 10, 2021. [Online]. Available: https://www.3gpp.org/technologies/keywords-acronyms/101-carrier-aggregation-explained.

[17] Nurul Huda Mahmood, Daniela Laselva, David Palacios, Mustafa Emara, Miltiades C. Filippou, Dong Min Kim, and Isabel de-la Bandera. Multi-Channel Access Solutions for 5G New Radio. *2019 IEEE Wireless Communications and Networking Conference Workshop (WCNCW)*, pages 1–6, 2019.

[18] TS 38.401: NG-RAN; Architecture description. V16.7.0, October 2021.

[19] Federico Lisi, Giacinto Losquadro, Andrea Tortorelli, Antonio Ornatelli, and M. Donsante. Multi-Connectivity in 5G terrestrial-Satellite Networks: the 5G-ALLSTAR Solution. *Ka and Broadband Communications, Navigation and Earth Observation Conference*, 2019.

[20] TS 22.261: Service requirements for the 5G system. V18.4.0, September 2021.

[21] Nurul Huda Mahmood, Melisa Lopez, Daniela Laselva, Klaus Pedersen, and Gilberto Berardinelli. Reliability Oriented Dual Connectivity for URLLC services in 5G New Radio. *2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–6, 2018.

[22] Hua Wang, Claudio Rosa, and Klaus I. Pedersen. Inter-eNB Flow Control for Heterogeneous Networks with Dual Connectivity. *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pages 1–5, 2015.

[23] Hua Wang, Guillermo Pocovi, Claudio Rosa, and Klaus I. Pedersen. Configuration of Dual Connectivity with Flow Control in a Realistic Urban Scenario. *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, pages 1–5, 2015.

[24] Nurul Huda Mahmood and Hirley Alves. Dynamic Multi-Connectivity Activation for Ultra-Reliable and Low-Latency Communication. *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, pages 112–116, 2019.

[25] Goksel Simsek, Hande Alemdar, and Ertan Onur. Multi-Connectivity Enabled User Association. *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–6, 2019.

[26] Valentin Poirot, Mårten Ericson, Mats Nordberg, and Karl Andersson. Energy efficient multi-connectivity algorithms for ultra-dense 5G networks. *Wireless Networks*, 26(3):2207–2222, Apr 2020.

[27] Michele Polese, Marco Giordani, Marco Mezzavilla, Sundeep Rangan, and Michele Zorzi. Improved Handover Through Dual Connectivity in 5G mmWave Mobile Networks. *IEEE Journal on Selected Areas in Communications*, 35(9):2069–2084, 2017.

[28] Marco Mezzavilla, Menglei Zhang, Michele Polese, Russell Ford, Sourjya Dutta, Sundeep Rangan, and Michele Zorzi. End-to-End Simulation of 5G mmWave Networks. *IEEE Communications Surveys Tutorials*, 20(3):2237–2263, 2018.

[29] Diomidis S. Michalopoulos, Andreas Maeder, and Niko Kolehmainen. 5G Multi-Connectivity with Non-Ideal Backhaul: Distributed vs Cloud-Based Architecture. *2018 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6, 2018.

[30] 5G ALLSTAR. Accessed on: Nov 3, 2021. [Online]. Available: https://5g-allstar.eu/.

[31] Deliverable D4.2 – Design and simulation of the multi-RAT load balancing algorithms. 5G ALLSTAR, 2019.

[32] F. Delli Priscoli, E. De Santis, A. Giuseppi, and A. Pietrabissa. Capacity-constrained Wardrop equilibria and application to multi-connectivity in 5G networks. *Journal of the Franklin Institute*, 358(17):9364–9384, 2021.

[33] Deliverable D4.3 – Implementation of the multi-RAT load balancing algorithms and technical specifications of the relevant interfaces. 5G ALLSTAR, 2021.

[34] E. De Santis et al. 5G-ALLSTAR Wireless Network Simulator. Accessed on: Dec 28, 2021. [Online]. Available: https://github.com/trunk96/wireless-network-simulator.

[35] Jocelyne Elias, Fabio Martignon, and Stefano Paris. Optimal Split Bearer Control and Resource Allocation for Multi-Connectivity in 5G New Radio. *2021 Joint European Conference on Networks and Communications 6G Summit (EuCNC/6G Summit)*, pages 187–192, 2021.

[36] Cristian Tatino, Ilaria Malanchini, Nikolaos Pappas, and Di Yuan. Maximum throughput scheduling for multi-connectivity in millimeter-wave networks. *2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1–6, 2018.

[37] Luis Diez, Andres Garcia-Saavedra, Víctor Valls, Xi Li, Xavier Costa-Perez, and Ramón Agüero. LaSR: A Supple Multi-Connectivity Scheduler for Multi-RAT OFDMA Systems. *IEEE Transactions on Mobile Computing*, 19(3):624–639, 2020.

[38] Jesús Burgueño, Isabel de la Bandera Cascales, David Palacios Campos, and Raquel Barco. Traffic Steering for eMBB in Multi-Connectivity Scenarios. *Electronics*, 9:2063, Dec 2020.

[39] ns-3 | a discrete-event network simulator for internet systems. Accessed on: Nov 29, 2021. [Online]. Available: https://www.nsnam.org/.

[40] RP-193234: Solutions for NR to support non-terrestrial networks (NTN), WID, 3GPP TSG RAN meeting 86. 2019.

[41] Jani Puttonen, Timo Nihtilä, Mika Innanen, Riku Järvinen, Janne Kurjenniemi, Vesa Hytönen, and Tommi Flink. Cloud Simulation Platform for Satellite and Terrestrial Network Simulations. *Ka and Broadband Communications Conference, Sorrento, Italy, September 30 - October 2, 2019.*

[42] Magister SimLab, Network Simulations as a Service. Accessed on: Nov 29, 2021. [Online]. Available: https://www.magister.fi/services/.

[43] 5G-LENA simulator. Accessed on: Nov 29, 2021. [Online]. Available: https://5g-lena.cttc.es/.

[44] LENA LTE-EPC Network Simulator. Accessed on: Nov 29, 2021. [Online]. Available: http://networks.cttc.es/mobile-networks/software-tools/lena/.

[45] TS 36.331: Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification. V13.7.1, October 2017.

Appendix 1. Users' throughputs in the heterogeneous simulation scenario.

Table A1.1. Downlink application call throughput (kbps) per UE, 2000 UDP packets of 100B sent to each UE.

NodeId	MC_Off	MC_On_A3_based	MC_On_A4_based
6	196.8	279.8	282.6
7	338.3	359.0	378.9
8	330.4	457.8	450.5
9	349.2	210.9	235.4
10	394.7	373.8	371.2
11	437.1	623.0	610.7
12	197.1	199.9	233.8
13	432.3	475.2	498.6
14	226.2	449.0	488.7
15	401.1	401.4	408.0
16	331.0	477.0	403.0
17	342.9	355.0	320.9
18	290.5	313.1	313.8
19	317.5	319.6	322.2
20	250.4	361.9	380.2
21	268.4	341.4	342.2
22	330.7	401.2	400.6
23	289.9	436.1	447.1
24	349.6	419.0	421.0
25	212.8	246.0	240.2