UNIVERSITAT POLITÈCNICA DE CATALUNYA

DOCTORAL THESIS

Satellite Integration in 5G: Contribution on Network Architectures and Traffic Engineering Solutions for Hybrid Satellite-Terrestrial Mobile Backhauling

Author:
Jesús Fabián Mendoza Montoya

Supervisor:
Dr. Ramon Ferrús Ferré

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

Mobile Communications Research Group
Department of Signal Theory and Communications

February 17, 2019

To my parents

Abstract

The recent technological advances in the satellite domain such as the use of High Throughput Satellites (HTS) with throughput rates that are magnitudes higher than with previous ones, or the use of large non- Geostationary Earth Orbit (GEO) satellites constellations, etc, are reducing the price per bit and enhancing the Quality of Service (QoS) metrics such as latency, etc., changing the way that the capacity is being brought to the market and making it more attractive for other services such as satellite broadband communications. These new capabilities coupled with the advantages offered by satellite communications such as the unique wide-scale geographical coverage, inherent broadcast/multicast capabilities and highly reliable connectivity, anticipate new opportunities for the integration of the satellite component into the 5G ecosystem. One of the most compelling scenarios is mobile backhauling, where satellite capacity can be used to complement the terrestrial backhauling infrastructure, not only in hard to reach areas, but also for more efficient traffic delivery to Radio Access Network (RAN) nodes, increased resiliency and better support for fast, temporary cell deployments and moving cells.

In this context, this thesis work focuses on achieving better satellite-terrestrial backhaul network integration through the development of Traffic Engineering (TE) strategies to manage in a better way the dynamically steerable satellite provisioned capacity. To do this, this thesis work first takes the steps in the definition of an architectural framework that enables a better satellite-terrestrial mobile backhaul network integration, managing the satellite capacity as a constituent part of a Software Defined Networking (SDN) -based TE for mobile backhaul network. Under this basis, this thesis work first proposes and assesses a model for the analysis of capacity and traffic management strategies for hybrid satellite-terrestrial mobile backhauling networks that rely on SDN for fine-grained traffic steering. The performance analysis is carried out in terms of capacity gains that can be achieved when the satellite backhaul capacity is used for traffic overflow, taking into account the placement of the satellite capacity at different traffic aggregation levels and considering a spatial correlation of the traffic demand. Later, the thesis work presents the development of SDN-based TE strategies and algorithms that exploits the dynamically steerable satellite capacity provisioned for resilience purposes to better utilize the satellite capacity by maximizing the network utility under both failure and non-failure conditions in some terrestrial links, under the consideration of elastic, inelastic and unicast and multicast traffic. The performance analysis is carried out in terms of global network utility, fairness and connexion rejection rates compared to non SDN-based TE applications.

Finally, sustained in the defined architectural framework designs, the thesis work presents an experimental Proof of Concept (PoC) and validation of a satellite-terrestrial backhaul links integration solution that builts upon SDN technologies for the realization of End-to-End (E2E) TE applications in mobile backhauling networks with a satellite component, assessing the feasibility of the proposed SDN-based integration solution under a practical laboratory setting that combines the use of commercial, experimentation-oriented and emulation equipment and software.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Dr. Ramon Ferrus, for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. I would also like to thank the National Council of Science and Technology (CONACYT) of Mexico for the financial support. I would like to give a special recognition to my parents for their great love, support and motivation throughout this journey. Likewise, I would like to give a special thank you to my daughter for her patience.

General Index

List of Figures

List of Tables

List of Abbreviations

2G	Second Generation
2.5G	2.5 Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
5G	Fifth Generation
A-CPI	Applications-Controller Plane Interface
A-DPI	Data-Controller Plane Interface
ACF	Autocorrelation Function
ACTN	Abstraction and Control of Transport Networks
API	Application Programming Interface
APP	Application
ARPU	Average Revenue Per User
ATM	Asynchronous Transfer Mode
BS	Base Station
BSM	Broadband Satellite Multimedia
C	Customer
CAL	Control Abstraction Layer
CAPEX	Capital Expenditures
CD	Coherence Distance
CDF	Cumulative Distribution Function
CIM	Common Information Model
CMI	CNC-MDSC Interface
CNC	Customer Network Controller
CP/C-plane	Control Plane
DAL	Device and Resource Abstraction Layer
DTH	Direct To Home

DVB-S	Digital Video Broadcasting by Satellite
E-UTRAN	Evolved Terrestrial Radio Access Network
E2E	End-to-End
eNB	E-UTRAN Node B or Evolved Node B
ETH	Ethereum
ETSI	European Telecommunications Standards Institute
FBS	Fixed Base Station
FCPAS	Fault, Configuration, Performance, Accounting and Security
GBR	Guaranteed Bit Rate
GEO	Geostationary Earth Orbit
GMPLS	Generalized Multi-Protocol Label Switching
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HTS	High Throughput Satellites
IETF	Internet Engineering Task Force
IM	Microwave Information Model
IMP	Information Modelling Project
IP	Internet Protocol
ISD	Inter-site Distance
ISO	International Organization for Standardization
LTE	Long Term Evolution
M&C	Management and Control
MAL	Management Abstraction Layer
MDSC	Multi-Domain Service Coordinator
MG	Multicast Guaranteed Bit Rate
MNO	Mobile Network Operator
MP/M-plane	Management Plane
MPI	MDSC-PNC Interface
MPLS	Multi-Protocol Label Switching
NBI	Northbound Interface
NCC	Network Control Centre

NE	Network Element
NETCONF	Network Configuration Protocol
NFV	Network Function Virtualization
NMC	Network Management Centre
Node B	3GPP term for a Base Station
NR	New Radio
NSAL	Network Services Abstraction Layer
NUM	Network Utility Maximization
OCH	Optical Chanel
ODU	Optical Data Unit
OF	OpenFlow
OF-Config	OpenFlow Management and Configuration Protocol
ONF	Open Networking Foundation
OPEX	Operating Expense
OSI	Open Systems Interconnection
OTN	Optical Transport Network
P-GW	Packet Data Network Gateway
PCE	Path Computation Element
PDF	Probability Density Function
PDN	Packet Data Network
PNC	Physical Network Controller
PoC	Proof of Concept
QIDs	Queue Identifiers
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RCS	Return Channel Satellite
RESTCONF	Representational State Transfer Configuration Protocol
RFCs	Request For Comments
RNC	Radio Network Controller
S-GW	Serving Gateway

S	Satellite
SAP	Service Access Point
SBI	Southbound Interface
SD	Satellite Dependent
SDN	Software Defined Networking
SES	Satellite Earth Stations and Systems
SI	Satellite Independent
SNF	Satellite Network Function
SNO	Satellite Network Operator
SNS	Satellite Network Switch
ST	Satellite Terminal
T-API	Transport API
T	Terrestrial
TBS	Transportable Base Station
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
TE	Traffic Engineering
TEAS	Traffic Engineering Architecture and Signalling
TP	Transport Profile
TR	Technical Report
TV	Television
U-plane	User Plane
UE	User Equipment
UG	Unicast Guaranteed Bit Rate
UGW	Gateway ST
UN	Unicast Non-Guaranteed Bit Rate
UST	User ST
WSON	Wavelength Switched Optical Network

Chapter 1

Introduction

1.1 Scope and Motivation

The recent development in wireless communications systems mainly pushed by the great boom of the cellular telephony, have dispersed its use to almost all sectors of industry and society. This trend is expected to continue over the next years, to the extent that 5G networks are envisioned to be used as the primary means for delivering applications with high availability needs in many sectors such as critical infrastructures, manufacturing, emergency communications, automotive, health, etc [1]-[3]. Given the relevance that mobile communications have taken in the recent years, 5G networks are expected to fulfill new challenging requirements such as higher network availability levels (at least in the range of the five nines, i.e. 99.999% of availability) [1], in addition to ubiquitous broadband connectivity extended to rural and low-density areas as well as long-haul transportation (e.g. aircraft, trains), among others [1][3].

In this context, the role that satellite communications can play in the forthcoming 5G ecosystem is being revisited [2]-[4]. During the past decades, satellite systems have suffered from many drawbacks such as technological complexity, high costs, high latency and deep fading at high frequencies (Ka band). As a result, for decades the use of satellite as means of communication had been mainly restricted to satellite niche market (e.g., professional use in areas where terrestrial alternatives are not an option, radio localization [GPS, GNSS], where satellites work as "radio-beacons" [5], Direct-To-Home TV market [DTH] digital TV broadcasting [standard DVB-S] where

the satellite acts only as a relay node, and more recently the backhauling of data in remote areas). All of them, applications where the use of satellite communication systems is usually driven by its intrinsic capabilities, for instance very large coverage, speed of implementations and inherent multicasting and broadcasting capabilities, exploiting these advantages in a competitive manner with respect to classical terrestrial networks [5]. However, recent technological advances in the satellite domain such as the introduction of High Throughput Satellites (HTS) in Geostationary Earth Orbit (GEO) with throughput rates that are magnitudes higher than with previous ones, with expectations of over 100 HTS systems in orbit by 2020-2025, delivering Tbps of connectivity in Ku- and Ka- bands at reduced cost [6], or the appearance of a range of disruptive initiatives envisioning the use of non-GEO constellations with a large number of low-cost micro-satellites that might come to fruition in the forthcoming years, anticipating a further reduction in the cost of transmitting a bit over a satellite link and enhanced Quality of Service (QoS) metrics such as latency, etc. [7], are changing the way that capacity is being brought to the market, reducing the price per bit and making it more attractive for other services such as satellite broadband communications. In this context, key features of satellite communications such as the unique wide-scale geographical coverage, inherent broadcast/multicast capabilities and highly reliable connectivity, together with significant amounts of new satellite capacity, opens an important range of possibilities for the implementation of satellite component into the terrestrial communication networks. For example, according to [8], non-terrestrial access networks are expected to be an integral part of 5G service deployment by:

- Enabling ubiquitous 5G service to terminals (especially Internet of Things [IoT] / Machine Type Communications [MTC], public safety/critical communications) by extending the reach of terrestrial based 5G networks to areas that cannot be optimally covered by terrestrial 5G network.

- Enabling 5G service reliability and resiliency due to reduced vulnerability of air / space borne vehicles to physical attacks and natural disasters. This is especially of interest to public safety or railway communication systems.

- Enabling connectivity of 5G Radio Access Network (5G-RAN) elements to allow ubiquitous deployment of 5G terrestrial network.
- Enabling connectivity and delivery of 5G services to User Equipment (UE) on board airborne vehicles (e.g. air flight passengers, UASs/drones, etc.).
- Enabling connectivity and delivery of 5G services to UE on board other moving platforms such as vessels and trains.
- Enabling efficient multicast/broadcast delivery of services such as A/V content, group communications, IoT broadcast services, software downloads (e.g. to connected cars) and emergency messaging.
- Enabling flexibility in traffic engineering of 5G services between terrestrial and non-terrestrial networks.

Remarkably, the potential roles and benefits of satellite networks in 5G have been acknowledged within the 3 rd Generation Partnership Project (3GPP) in charge of 5G system specifications, where requirements have been adopted that mandate 5G systems to be able to provide services using satellite access as well as to support the use of satellite backhaul between the Radio Access Network (RAN) and core network by enhancing the 3GPP system to handle the latencies introduced by satellite links [9]. Accordingly, several 3GPP study items are currently on-going to undertake the necessary technical studies for the support of satellite as well as High Altitude Platforms (HAPS) as part of next generation 5G systems [10].

In this context, satellite backhaul stands out as one of the most compelling scenarios anticipated to gain further relevance in 5G [11]. In particular, satellite backhauling can be instrumental in providing backhaul connectivity to base stations (BSs) deployed in hard to reach areas or installed on board a transportation/moving vehicle with no other feasible backhauling means. In addition, satellite backhaul links can be deployed and operated in combination with the terrestrial backhaul links, yielding increased network availability and resiliency (e.g. backup capacity for total/partial terrestrial link failures in critical cell sites), better support of temporary cell deployments (e.g. coverage of special events, emergency situations) [12] and, ultimately, more efficient traffic delivery to BSs. For example, jointly exploited

with the terrestrial capacity, a pool of satellite capacity can be used for traffic offloading and load balancing (e.g. diverting traffic from congested areas so the terrestrial capacity gets supplemented during peak-times) as well as for multicast/broadcast traffic delivery to multiple cell sites (e.g. content edge caches, live TV stream distribution) in a more resource efficient manner [12][13].

However, a full realization of a combined satellite-terrestrial backhauling scenario requires of improved integration approaches for the provisioning and operation of the satellite component in a more flexible, agile and cost-effective manner than currently done. In this regard, the evolution of satellite ground segment systems (e.g. satellite gateways and terminals) from today's rather closed solutions towards more open architectures based on Software Defined Networking (SDN) and Network Function Virtualization (NFV) technologies arises as a necessary step [14]-[18]. This evolution is not only expected to bring into the satellite domain the benefits associated with the advances in network softwarization technologies consolidating within the 5G landscape, but also to greatly facilitate the seamless integration and operation of combined satellite and terrestrial networks [19]. In particular, terrestrial 5G systems are widely embracing SDN technologies for enabling a unified, vendor-neutral control and management of networking functions. Accordingly, satellite networks shall be necessarily outfitted with control and management interfaces compatible with these mainstream SDN architectures and technologies being adopted in 5G in order to realize a full End-to-End (E2E) networking concept where the whole satellite-terrestrial network behavior can be programmed in a consistent and interoperable manner [20]-[22].

In the other hand, sustained into the introduction of SDN technologies for the integration of the satellite component, Traffic Engineering solutions adapted to this kind of networks should be also revisited and developed, taking all the advantages of a centralized control and fine-grained traffic steering offered by SDN networks in order to optimize the performance of a hybrid mobile backhaul network, taking into account the different performances and capacities of satellite and terrestrial networks as well as the different QoS requirements of the mobile network services. In this sense, despite the multiple advantages mentioned above, the reality as of today

is that the combination of satellite and terrestrial components for delivery of end-to-end telecom services as well as the development of TE solutions for the optimal traffic distribution for this kind of networks still remains as completely unexplored areas. While capacity and traffic management aspects (e.g., capacity dimensioning, routing, QoS, congestion, resilience) have been extensively analysed for mobile terrestrial backhaul networks [23]-[27], little attention has been paid to these aspects when it comes to the consideration of hybrid terrestrial-satellite networks for mobile backhauling. In this line, there is a wide range of possible contributions to the subject, allowing this research to be one of the first steps in the development of new studies of the aggregate capacity gain by the introduction of given satellite capacity into a terrestrial mobile backhaul networks as well as the development of traffic distribution algorithms focused on maximizing the utility of this kind of networks, considering different kind of traffics (elastic, inelastic, unicast and multicast traffic) as well as considering the differences in capacities and performance of both kind of networks (satellite and terrestrial).

In the following, the objectives and contributions proposed for the development of this thesis are explained in detail.

1.2 Objectives and Contributions

The following points break down the main objectives pursued by the thesis:

1. To progress beyond the state-of-the-art by investigating in depth in the following points:

    - The applicability of the SDN-based paradigm to the combined satellite-terrestrial networking.
    - The architectural frameworks based on SDN controllers and federated resource managers for the orchestration and optimization of network functions for the delivery of services that span satellite and terrestrial segments in order to facilitate and improve the combination of satellite and terrestrial networks in the context of 5G.

  - Specifically in the context of satellite backhaul, in the definition of SDN-based management principles to create new degree of freedom to ensure smooth interoperations among the fronthaul, backhaul, the core network layers, while separating the bearer & control functions via SDN and centrally managing.

- The design and analysis of current models for the analysis of capacity and traffic management strategies for hybrid satellite-terrestrial mobile backhaul networks.
- The Traffic distribution strategies (based on QoS, load balancing, network utility applied for current terrestrial networks and for hybrid satellite-terrestrial mobile backhaul networks.

2. Based on the state of the art research work, to define an architectural framework that better facilitates the integration of the satellite and terrestrial components to form a hybrid satellite-terrestrial mobile backhaul network.

3. On top of a defined architectural framework, to analyze the hybrid satellite-terrestrial mobile backhaul networks:

   - A first proposal of a model for the analysis of capacity and traffic management strategies in hybrid satellite-terrestrial mobile backhauling networks that rely on SDN for fine-grained traffic steering.
   - Simulation activities for the capacity assessment for this type of networks in realistic environments.

4. On top of a defined architectural framework, to design TE applications for hybrid satellite-terrestrial mobile backhaul networks.

   - The work tries to integrate and exploit satellite features in an SDN perspective, searching to propose new TE for traffic distribution strategies and algorithms for joint capacity and traffic management across terrestrial and satellite backhaul network elements with the following considerations:

- To better exploiting the dynamically the steerable satellite capacity in hybrid satellite-terrestrial backhaul networks provisioned for resilience purposes in order to maximize a network utility function.
- Management of provisioned resources (e.g., radio and spectrum resource allocation), QoS classes and policy control across domains is also under scope, as well as the exploitation of off-loading mechanisms.
- Under the consideration of different kinds of traffic types (elastic, inelastic, unicast and multicast).
  * Consideration of QoS of each traffic Type
  * Consideration of satellite and terrestrial performance (latency) impact over each traffic type.
- Traffic distribution strategies for resilience purposes.
  * Adopting mechanism of resource allocation and resource reservation when some Base Stations (BSs) have terrestrial link failures.
  * Satellite links could provide additional bandwidth to backup connectivity to critical cell sites as well as to divert traffic from congested areas so that the capacity in the terrestrial links could be supplemented during peak-times or even replaced in case of total/partial failure or maintenance.

5. The thesis work search to evaluate all the proposed solutions through assessment activities that are intended to be based on the use of simulators/emulators tools that are developed and evaluated to know better the behavior of the hybrid networks in real scenarios and conditions where the proposed mechanisms and algorithms are applied. This with the objective of evaluating the resource management strategies and algorithms for the combined operation of the terrestrial and satellite components, as well as the assessment of the implementation of the satellite component as an intrinsic part of terrestrial communications networks under different scenarios and conditions, analyzing the scope, advantages, disadvantages and viability of such combined networks.

6. The final objective of this thesis work is to present an experimental Proof of Concept (PoC) of a satellite-terrestrial backhaul links integration solution that builts upon SDN technologies for the realization of End-to-End (E2E) TE applications in mobile backhauling networks with a satellite component, with the objective of know the scope, advantages and disadvantages and applicability of these kind of networks as well as the performance impact that it has on the QoS of the LTE connectivity services.

1.3 Organization of the Thesis

The thesis is divided into 8 chapters as follows:

In chapter 2, **Hybrid Satellite-Terrestrial Mobile Backhaul Networks**, a brief description of basic concepts related to Hybrid Satellite-Terrestrial Mobile Backhaul Networks such as satellite communication systems, mobile backhaul networks and network architectures is presented, as well as a review of the role of satellite technology in 5G networks. The knowledge of these basic issues will set the tone for the further developments of the specific objectives of the research work, together with the methodology to be used along in the research activities.

In chapter 3, **Network Architecture and Integration Approach for SDN-based Traffic Engineering in Satellite-Terrestrial Backhaul Networks**, is defined the SDN-based functional architecture for hybrid satellite-terrestrial mobile backhaul networks on which the capacity analysis and the development of traffic distribution strategies for this kind of networks will be carried out. Also it's presented an integration approach for E2E SDN-Based TE in hybrid satellite-terrestrial mobile backhaul networks.

In Chapter 4, **Backhaul capacity gains of traffic overflow through satellite**, a dimensioning study of the satellite backhaul capacity required at BS and in a higher aggregated level is carried out. The analysis provided is aimed to assess the capacity

gains when a satellite component is used for traffic overflow within the hybrid backhaul network, taking into account the placement of the satellite capacity at different traffic aggregation levels considering the spatial correlation of the traffic demand. The results of the analysis show substantial improvements in terms of capacity gains when the satellite backhaul capacity is used for traffic overflow, at different traffic aggregation levels.

In chapter 5, **Network Utility Maximization Framework and Performance Evaluation of TE strategies**, is presented the Network Utility Maximization (NUM) Framework, and are proposed TE strategies for optimal traffic distribution (maximizing the network utility) in SDN-based hybrid satellite-terrestrial mobile backhaul networks considering elastic and inelastic traffic types. A performance analysis is carried out, finding substantial improvements in terms of network utility, fairness and services rejection rates. The methodology utilized to solve the optimization problem is a Heuristics and is assessed through Monte-Carlo simulations.

In chapter 6, **Formulation and performance assessment of an SDN-based TE application**, it is formulated a TE application that takes advantage of the SDN-based centralized control to manage traffic in a better way in the presence of both, failure or non-failure terrestrial link conditions. In particular, the focus has been placed in the investigation of SDN-based the TE application designed to deal, in a comprehensive manner, with situations that entail traffic overload in some BS sites, fast or early stage deployment of RAN capacity (i.e. transportable BSs) that exclusively relies on the satellite component and total or partial failure of the terrestrial links in BSs. The simulation model developed in the previous chapter has been enriched with the consideration of multi-service scenarios, including mixes of stream and elastic traffic as well as unicast and multicast traffic. Moreover, the simulation platform used in the previous chapter analysis, which was based on Monte-Carlo simulation methods, has been extended to support discrete event simulations to track the system dynamics over time, allowing us to improve the characterization of the SDN-based TE strategies with more practical settings (e.g. time average processes, timer-based triggering conditions, etc.) as well as to obtain some performance indicators that cannot

be assessed with the Monte-Carlo method. In more details, the analyzed SDN-based TE application is formulated to manage the use of some amount of satellite capacity provisioned for resilience purposes so that the overall network utility is maximized under both failure and non-failure conditions in the terrestrial links. The new algorithm considers the analysis for multicast traffic and introduces the mechanism for the reserve of satellite resources for the BSs that only have satellite capacity. Several event-driven simulations and scenarios are presented to perform a complete algorithms evaluation under mobility and normal and failure terrestrial link conditions.

In chapter 7, **Experimental Proof of Concept**, is presented an experimental Proof of Concept (PoC) and validation of a hybrid satellite-terrestrial mobile backhaul network integration solution that builts upon SDN technologies for the realization of End-to-End (E2E) TE applications in mobile backhauling networks with a satellite component. A laboratory testbed is developed and validated, then, it is assessed the LTE service performance over terrestrial and satellite backhaul links. Finally, it is assessed the network the backhaul links restoration time once a given backhaul link failure occurs and its QoS impact on LTE services.

In chapter 8, **Conclusions and Contributions**, are listed the final conclusions and contributions of this thesis work.

Appendix -A- Description of the main features and building blocks of the matlab simulator developed for Network Utility Maximization through Monte-Carlo simulations;

Appendix -B- Description of the main features and building blocks and Matlab simulator for TE application performance assessment.

The structure of the thesis is illustrated in Figure 1.1.

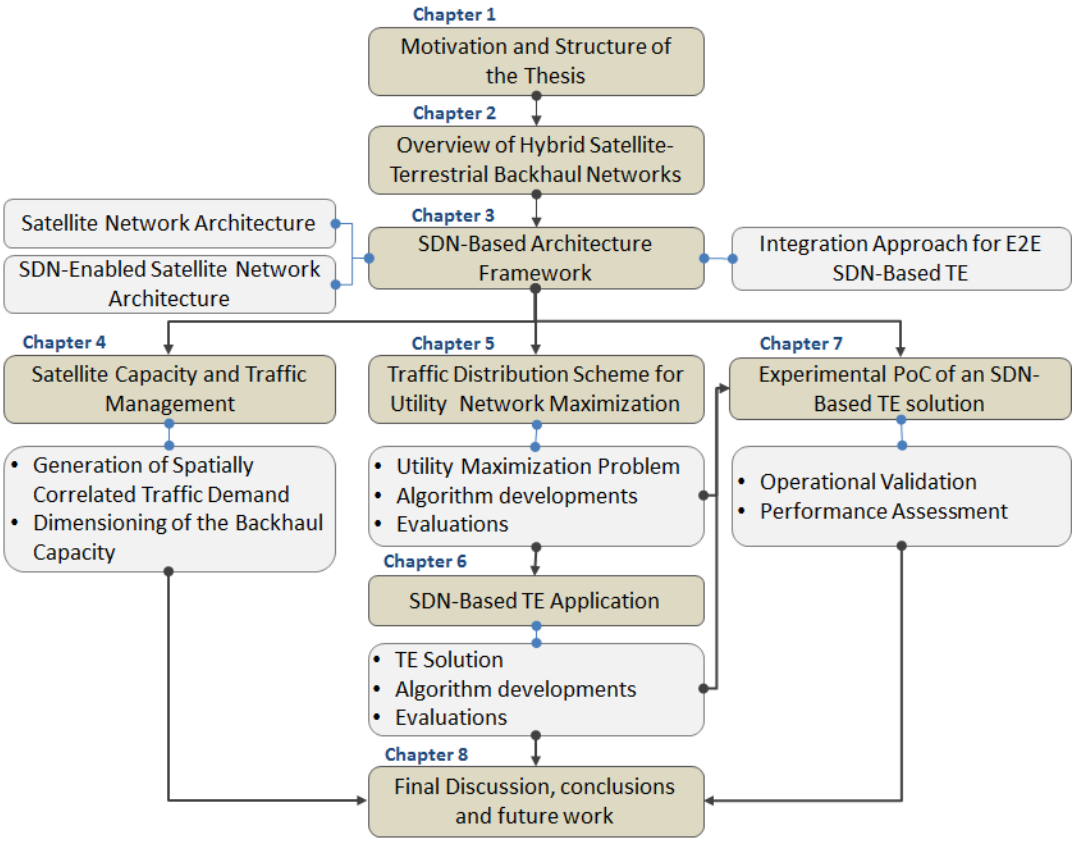

FIGURE 1.1: Structure of the Thesis.

1.4 Main Outcomes

1.4.1 Publications

Part of the content of this thesis has been either been published or submitted for publication during the period of research of the author in the Mobile Communication Research Group of the Department of Signal Theory and Communications, at Universitat Politècnica de Catalunya. A list of the papers is given as follows:

- F. Mendoza, R. Ferrús, O. Sallent, *Experimental Proof of Concept of an SDN-based Traffic Engineering Solution for Hybrid Satellite-Terrestrial Mobile Backhauling*. Submitted to International Journal of Satellite Communications and Networking. Oct. 2018. Article accepted on 10 February, 2019.

- F. Mendoza, R. Ferrús, O. Sallent, Chapter 3 , *SDN-enabled SatCom Networks for Satellite-Terrestrial Integration* of book *Satellite Communications in the 5G Era*, IET Digital Library, ISBN: 978-1-78561-427-9, 2018. pp 61-99.

- F. Mendoza, R. Ferrús, O. Sallent, *SDN-based Traffic Engineering for Improved Resilience in Integrated Satellite-Terrestrial Backhaul Networks*, 4th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Münster, Germany. Dec. 2017.

- F. Mendoza, R. Ferrús, O. Sallent, *A traffic distribution scheme for 5G resilient backhauling using integrated satellite networks*, 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain. June. 2017.

- F. Mendoza, R. Ferrús, O. Sallent, *Flexible Capacity and Traffic Management for Hybrid Satellite-Terrestrial Mobile Backhauling Networks*, International Symposium on Wireless Communication Systems (ISWCS), Poznan, Poland. Sept. 2016.

1.4.2 Contributions to European Research Project VITAL Deliverables

Part of the work done in the scope of the thesis has served as a basis for the realization of some research activities that have been disseminated in technical reports, participating as contributor in the following research projects deliverables:

- D4.2 *Network resource management framework and initial performance assessment of algorithmic solutions*, Editor: Ramon Ferrús, August 2016. Available at `http://www.ict-vital.eu/documents/deliverables`. Last accessed in December 2018.

- D4.5 *Final network resource management framework and performance assessment of algorithmic solutions*, Editor: Ramon Ferrús, August 2018. Available at `http://www.ict-vital.eu/documents/deliverables`. Last accessed in December 2018.

1.4.3 Simulation Tools and Experimental Platforms

As part of the results of this thesis, simulation tools have also been developed, among which the following unravel:

- Simulation of spatial correlation of the traffic demand as well as the capacity gains for satellite capacity at different traffic aggregation levels in a hybrid backhaul network.

- Matlab simulator for Network Utility Maximization through Monte-Carlo simulations.

- Matlab simulator for TE application performance assessment.

- Experimental platform for TE application proof of concept.

Chapter 2

Hybrid Satellite-Terrestrial Mobile Backhaul Networks

2.1 Introduction

A general view on systems combining satellite and terrestrial communications components has been recently described in ETSI TR 103124 "Satellite Earth Stations and Systems (SES); Combined Satellite and Terrestrial Scenarios" [28] where the definitions and classification of scenarios combining satellite networks as well as terrestrial networks are proposed. These scenarios are showed at Figure 2.1.

A "combined satellite-terrestrial network" is a system employing a satellite component and a terrestrial component to deliver a service set towards its end-users/subscribers [28]. Both components may be controlled by the same network management system and possibly use the same portions of frequency band allocation. The satellite component may operate in parallel to the terrestrial component or may operate as backhaul to the terrestrial component from the end-user terminal point of view.

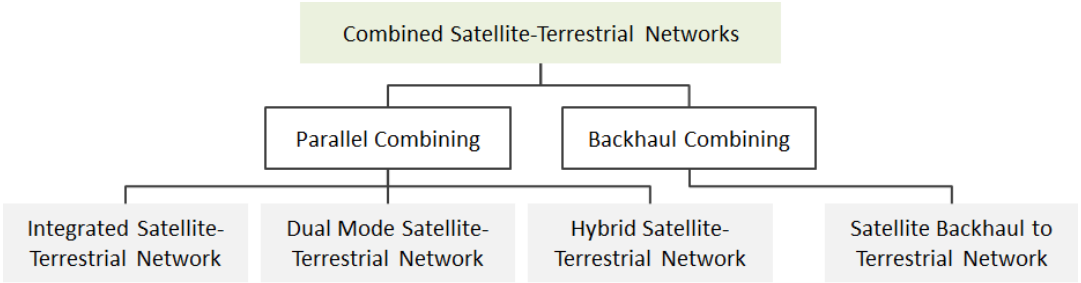

FIGURE 2.1: Combining Satellite-Terrestrial Scenarios, Proposed by [28].

Satellite backhaul stands out as one of the most compelling scenarios anticipated to gain further relevance in 5G [11]. Then, a Hybrid Satellite-Terrestrial Backhaul Network is then characterized by the capability of delivering a backhaul service using simultaneously a satellite and terrestrial components. Under this scenario, the satellite backhaul network is used in both directions to provide bulk connectivity to a terrestrial network element (e.g. to a cellular base station or to a local area network, etc.).

In the rest of the chapter, a brief description of basic concepts related to Satellite Communication Systems and Mobile Backhaul Networks is presented. The concepts and definitions of this chapter will set a basis to present the further developments and algorithms proposed by the thesis in later chapters.

2.2 Limitations of Current Platforms

Due to much lower economies of scale and inherent associated technological complexities, satellite communications offerings have not evolved at the same pace as terrestrial communications systems have done, particularly when compared to mobile broadband communications. As a matter of fact, out of the direct-to-home TV market along with satellite niche areas as coverage of air and sea, satellite communications are still considered neither a flexible nor a cost-effective solution in most of the cases and are just exploited as a last resort or worst case transport solution. This is mainly due to certain key limitations of current platforms. In the following are explained the limitations of current platforms that mainly define the problem to be solved for the implementation of the hybrid satellite-terrestrial mobile backhaul networks, as well as the technological advances and trends in the satellite domain that motivate the development and provide viability to this kind of networks.

Some key limitations of current platforms are the following [19]:

- The establishment and configuration of networking services across satellite and terrestrial domains is mostly performed manually, thus involving considerable setup and reconfiguration delays, as well as high associated operating and maintenance costs.

- The new network technologies, algorithms and protocols can't be rapidly introduced into the market since they involve time-consuming and costly hardware upgrades that are associated with significant CAPEX investments.
- The satellite resources are assigned statically to each user, without the capability of automatically up/down the scale accordingly to users demand.
- Satellite communication services are mostly associated with plain connectivity (without QoS), without the ability to insert on-demand in-network services (e.g., firewalling, proxying for traffic optimization, caching, media transcoding, etc.) for network-side traffic processing.
- There are many satellite specific settings and the lack of common prevalent standards for their integration with terrestrial systems do not provide a transparent manner for the applicability and continuity of policies for e.g., routing, quality of service, security, management and connectivity (Ethernet, MPLS, etc.), across both segments.

2.3 Technological Advances and Trends in the Satellite Domain

The use of satellite links in terrestrial networks is an issue increasingly addressed by different authors and more and more companies are opting for them. This trend emerges according to various triggers and some industry trends that have achieved the viability of download and overflow networks through satellite links. Some triggers and trends are:

- *The development of new High throughput satellite platforms (HTS)*.- The new generation of high throughput satellites (HTS) with the largest number of spots can deliver a capacity of more than 90 Gbit/s, while VHTS systems (e.g., Viasat-3) aim at achieving data rates in the range of Tbps in Ku- and Ka- bands at reduced cost, with expectations of over 100 HTS systems in orbit by 2020-2025 [6][29].

- *The new HTS Multi-beam capacity*.- With the new HTS platforms, the operators can provision a bandwidth that can be dynamically assigned through different beams, based on the demand. This extends the HTS economy to wider geographical areas and eliminates the need of purchasing a fixed amount of capacity per beam [6][30].

- *Higher Flexibility*.- The satellite technology has historically been benefited during disruptions of traffic demands because its indifference to the distances, and it can flexibly support bandwidth needs. Even in areas with high population density, it takes time to solve problems related to traffic congestion on backhaul links. But one based on satellite equipment solution can quickly (in less than one day) be installed in the base stations. This improves the user experience (immediately eliminates congestion) while the operator gets time for engineering and deploying a terrestrial backhaul solution.

- *Introduction of Satellite Terminals with lower cost and higher speed*.- This is facilitated by the high volume market consumers. HTS has reached a point of scale that allows high-speed connectivity at low CAPEX. The satellite terminals cost hundreds of dollars, which is negligible in the context of the costs of some microwave transmission equipment and within the range of cost of femto-cell deployments.

- *Lower Cost-Benefit ratio*.- Increasing the capacity of terrestrial backhaul links typically involves the addition of communications equipment in multiple locations (endpoints, repeaters, etc.), increasing the CAPEX. However, one satellite terminal can absorb traffic and can differentiate intelligently the route traffic to the satellite and the traffic rerouted by the terrestrial link. It is anticipated that there will be a set of scenarios where the satellite OPEX will be amply justified by the savings in CAPEX.

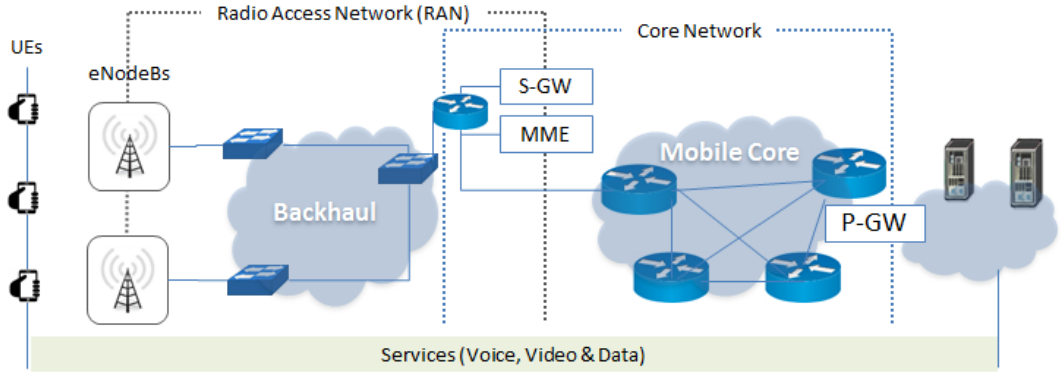

FIGURE 2.2: Typical Mobile Network Deployment.

2.4 Mobile Backhaul Networks

The term backhaul network or simply backhaul are commonly referred as the part of the network comprises the intermediate links between the core or backbone network and the small sub networks at the "edge" of the entire hierarchical network. In mobile networks, the backhaul networks is commonly referred to the part of mobile access network that interconnects the network elements (NE) in the RAN to the core network by means of different topological configurations and transmission technologies. Therefore, in the LTE systems the backhaul interconnects the eNB to the S-GW, which are subsequently connected to data centers that host the content and applications accessed by mobile users—both human and machines. The Figure 2.2 presents schematically an example of a typical mobile network deployment where the function of the backhaul network is clearly shown.

As mobile standards usually provide some flexibility in how the backhaul should be implemented, mobile operators can adopt different strategies and rely on available technologies in the market to build its own backhaul infrastructure or also lease part of the transmission network to a carrier. Focusing on the backbone part of the mobile backhaul network, the transmission infrastructure used in this stage is normally optical fiber since high volumes of traffic are consolidated at each of the connection points, however, the transmission infrastructure in the backhaul network (i.e., last mile and second mile stage) is more difficult to deploy. This is because there are different factors that influence the selection of the physical medium to be deployed in these two stages to interconnect the different elements in the backhaul.

This may include the BS density, terrain characteristics and distances, the target coverage area (urban or rural), required capacity, the availability of transmission technologies and finally a very important one, the economic factor. These factors influence in an important way the choice of technologies and design of the backhaul network. For example, nowadays the mobile network operators face important challenges and constraints as a highly competitive and rapidly changing marketplace and a growing demand for data by the owners of smartphones and other devices, e.g., according to [31], between 2013 and 2014 there was a 60% growth in data traffic. Then, while it is necessary to constantly update the network backhaul capacity and mobile network coverage that involves high infrastructure deployments and searching to improve service quality, it is also necessary to increase profitability, even when there is a tendency to offer increasingly cheaper the bit per second. This leads to operators to seek higher average revenue per user (ARPU), taking major steps in introducing new technologies to increase their capacity but reducing the CAPEX and OPEX in mobile networks. In this aspect, the relevance that the backhaul network takes is also important due to the backhaul network is one of the mayor contributors to the high cost of building out and running a cellular network. In fact, a key challenge to mobile operators is to reduce backhaul costs. In this sense, mobile operators are continuously seeking for cost-effective solutions for the backhaul network in order to squeeze more out from the available network resources. While the mobile network backhaul increasingly becomes more relevant, as it is a component whose cost represents a significant share of CAPEX and OPEX of total mobile network, while allows to operators the deployment of new high capacity technologies as LTE-advanced.

2.5 LTE Backhauling over Satellite Systems

As mentioned before, the mobile networks are facing a surge in data traffic as smartphone use increases, devices become more affordable and network capacity expands. One way in which the technology are responding to this data explosion is by adopting new, small-cell architectures that can better target underserved areas in developed markets, and extend services into previously unconnected remote and rural

areas. A key challenge when servicing these small cells is the provision of backhaul. Fixed-line backhaul options, such as copper or fiber are impractical, inflexible or too costly for use in conjunction with small cells outside highly-developed areas. Due to its ubiquitous coverage, satellite technology is suitable for backhaul application in remote or rural areas where the market still does not justify the costs of deployment of other backhaul technologies.

Satellite technology is perfectly suited to LTE systems, and has already proved for mobile backhaul purposes. Modern TDMA-based systems can rapidly deliver the highly-cost-effective, carrier-class, two-way all-IP connectivity with shared bandwidth that is ideally suited to small-cell environments. The new satellite technologies offer IP connectivity at almost anywhere in the world. However, the challenge of this technology remains in the cost of the bandwidth of the link, which is still considerable even with the introduction of high-performance satellites, however, the links can be justified under certain scenarios, primarily as backhaul network.

Since satellite solutions can be set up quickly, communications networks and new services can be quickly recovered and reconfigured, reliable service and high availability, suitable for medium capacity backhaul links. In addition, you can expand services electronically without traditional terrestrial networks. As a result, you can achieve a high level of communications rapidly without high budget expenditures. Also, ideal for network deployments at inaccessible sites or as a complement to the previously installed infrastructure.

Under the backhauling scenario, the satellite network can operate in both forward and return direction to provide connectivity to a ground-based network component. The ground based component may be either a fixed platform (buildings or masts) or may be a moving platform (ships, trains, airplanes or other vehicles). Nowadays, several solutions are commercialized for 2G and 3G networks and new platforms and services are also starting to be offered using latest HTS satellites in Ka-band technology. The trend is that mobile backhaul is migrating to IP due to the older ATM and TDM-based technologies are becoming increasingly uneconomical as mobile broadband increases worldwide. IP-based satellite backhaul for remote and rural areas is being considered by mobile operators in emerging markets as a way to reduce costs while offering new opportunities for organic growth past the

already penetrated urban areas.

Satellite components integration with LTE terrestrial networks maximizes the likelihood of services take-up through identifying scenarios where satellites could enhance terrestrial networks in cost efficient manner [13]. In the next section, the possible scenarios and use cases identified by the research work for the application of a hybrid satellite-terrestrial mobile backhaul network are listed.

2.6 Hybrid Satellite-Terrestrial Mobile Backhaul Network Application Scenarios

Satellite component can be utilized as backhaul under three different scenarios of mobile communications as urban, rural or moving cells. In this line, we have made a classification of use cases where a hybrid satellite-terrestrial mobile backhaul network increases the backhaul capacity in different ways, which are shown illustratively in the Figure 2.3:

- *Satellite for traffic overflow*.- Satellite links may be used as traffic offload, ie, when a terrestrial backhaul link exceeds a threshold of use, close to congestion levels, the traffic excess may be routed through the satellite backhaul link, thus avoiding congestion the terrestrial link favoring the network performance for the end user.

- *Satellite for traffic offloading*.- The satellite links can be used as traffic offload of terrestrial backhaul, i.e., the satellite backhaul link can be used as a complement to the terrestrial backhaul, even when the terrestrial backhaul links have not exceeded the utilization thresholds, this in order to improve network performance to benefit the end user. Some approaches to the use of such links with the above purpose are as follows:

  - *To Increase the capacity of backhaul links*.- The use of satellite links as supplement the network of terrestrial backhaul must be based on a cost benefit study, besides considering other technical factors such as network topology, link capacity of satellite and terrestrial backhaul, latency, etc.

  - *For more efficient traffic delivery to RAN nodes*. More challenging is the use of the satellite component for smarter traffic offloading and load balancing strategies across the terrestrial and satellite backhauling components. For example, the satellite link can be used to offload multicast/broadcast traffic addressed to multiple cell sites (e.g. cached content at the RAN, TV live streams for onward multicast over the RAN) in a more resource efficient manner [13].
  - *Satellite for increased resilience*.- The satellite links can be used to increase the availability and resiliency of the mobile backhaul network. Satellite service can provide additional bandwidth to divert traffic from congested areas and backup connectivity to critical cell sites so that a limited capacity in their terrestrial links can be supplemented during peak-time or even replaced in case of total/partial failure or maintenance [12][32].

- *Satellite for moving cells*.- The use of small cells in a variety of mobile situations can be achieved simply by using satellite backhaul. These use cases include the use of small cells on terrestrial vehicles (e.g. trains, buses), aircrafts, and vessels [32][33]. For ships and airplanes special stabilized antenna systems are used to point at the satellite. The deployment of small cells inside transportation facilities (e.g. small cells within buses, trains, airplanes) is also a compelling case that deserves a flexible management of the backhauling capacity, which can be partly or fully reliant on satellite communications.

- *Satellite as primary backhaul for fast, temporary cell deployments*.- Small cells are compact and lightweight and are well suited to deployments of a temporary nature such as disaster recovery, first responder, and special events. Temporary deployments require equipment that is portable and that can be rapidly installed and commissioned to provide or restore essential communications infrastructure for special events or disaster recovery [11][32].

- *Satellite for rural and remote small cells*.- Small cell technologies are coming of age thanks to scaling of deployments in residential, enterprise and now urban markets. These technologies can now be applied to a range of rural and remote use cases due to a great availability of satellite links that otherwise may not

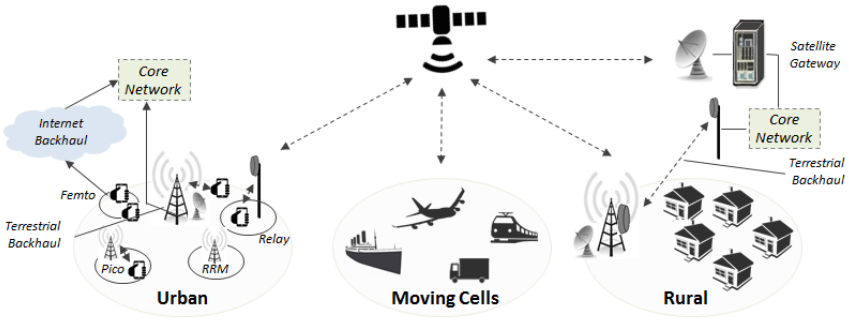

FIGURE 2.3: Hybrid Satellite-Terrestrial Mobile Backhaul Network Application Scenarios.

be viable using traditional deployment approaches. Low cost satellite backhauling together with advances in small cells create an appealing cost-efficient proposition for MNOs to extend their Radio Access Network (RAN) in areas underserved by terrestrial networks (e.g., rural and remote areas) or simply not possible to serve (e.g. maritime/aviation services) [33]-[36].

2.7 Summary

This chapter has presented a brief description of the Satellite communication systems and some concepts related to the combination of satellite systems with terrestrial systems to conform hybrid backhaul networks, such as the role of satellite technology in such hybrid networks and the identified application scenarios for hybrid satellite-terrestrial mobile backhaul networks.

Taking the concepts presented in this chapter, the next chapter explains the defined framework arquitecture aimed to greatly facilitate the seamless integration and operation of combined satellite and terrestrial mobile backhaul networks.

Chapter 3

Network Architecture and Integration Approach for SDN-based Traffic Engineering in Satellite-Terrestrial Backhaul Networks

3.1 Introduction

During the last decade the networking community is witnessing a paradigm shift towards more open architectures based on Software Defined Networking (SDN) and the softwarisation of communication networks in a quest for improved agility and flexibility, and ultimately cost reduction, in the deployment and operation of networks. In particular, terrestrial 5G systems are widely embracing SDN technologies for enabling a unified, vendor-neutral control and management of networking functions, decoupling the control plane from the user plane. In this regard, the so called Control and User Plane Separation (CUPS) architecture has been developed as an enhancement of the 4G/LTE standards to fully split control and user plane functions within the Evolved Packet Core (EPC) [37]. Likewise, the new 5G Core Network (5CN) specifications have consolidated this separation as a key design principle [38]. In this line, satellite networks shall be outfitted with a set of control and management functions and interfaces (API and/or network protocols) compatible

with the mainstream SDN architectures and technologies being adopted in 5G in order to realize a full End-to-End (E2E) networking concept where the whole satellite-terrestrial network behavior can be programmed in a consistent and interoperable manner. To achieve a high degree of radio interface commonality and higher layer operational integration with 5G terrestrial access, the deployment and operation of networks that combine terrestrial and satellite transmission components is also expected to benefit from the incorporation of network softwarisation technologies such as Software Defined Networking (SDN) and Network Function Virtualization (NFV) [14][18][19][39] into satellite systems.

In this line, this chapter defines and explains a SDN-based functional architecture for the ground segment of a satellite broadband communications system, delineating the different alternatives for supporting SDN concepts and technologies within both internal and external interfaces of satellite networks. In order to set the ground for the discussion, key foundations on SDN architectures and technologies as well as on satellite broadband system architectures are briefly outlined first. Likewise, it is presented an integration approach for E2E SDN-Based TE in hybrid satellite-terrestrial mobile backhaul networks.

3.2 SDN-Based Functional Architectures for Satellite Networks

General principles and reference SDN architectures have been specified by the Open Networking Foundation (ONF) and Internet Engineering Task Force (IETF) in [40] and [41], respectively. Both SDN architectural models are illustrated in 3.1 Keeping aside some differences in terminology and orientation, both architectures reflect the key principles of SDN: (1) separation of data plane resources (e.g. data forwarding functions) from control and management functions; (2) centralization of the management-control functions, and (3) programmability of network functionality through device-neutral and vendor-neutral abstractions and APIs. While the IETF model description is more centered on network devices and control and management abstraction layers, the ONF model is specified around the so-called SDN controller, which is the core functional entity of the SDN architecture. The SDN controller exposes services and resources to clients via applications-controller

plane interfaces (A-CPIs), and consumes underlying services and resources via data-controller plane interfaces (D-CPIs). A-CPIs and D-CPIs are, respectively, the equivalents of the Control Plane / Management Plane Southbound Interfaces (CP/MP SBIs) and Service Interfaces within the IETF model. Service Interfaces are also commonly referred to as Northbound Interfaces (NBI).

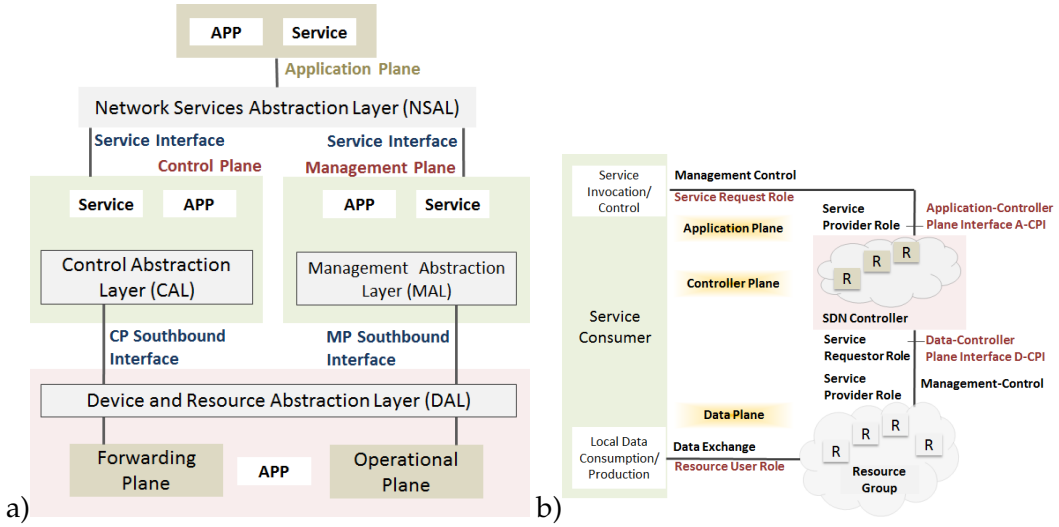

FIGURE 3.1: IETF RFC 7426 and ONF SDN Architectural Models;(a) IETF RFC 7426 SDN Architectural Model [41], (b) ONF SDN Architectural Model [40].

A more purpose specific SDN architecture for transport networks is being developed by the Traffic Engineering Architecture and Signaling (TEAS) Working Group within the IETF, which is responsible for defining Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) traffic engineering architectures and protocols. Such SDN architecture, named Abstraction and Control of Transport Networks (ACTN), describes a control framework for operating a TE network (such as an MPLS-TE network or a layer 1 transport network) to provide connectivity and virtual network services for customers of the TE network. The services provided by the ACTN can be tuned to meet the requirements (such as traffic patterns, quality, and reliability) of the applications hosted by the customers. An illustration of the ACTN architecture is given in Figure 3.2. The ACTN architecture is well aligned with the previously introduced ONF and IETF SDN architectural principles even through it is represented as a 3-tier reference model. Importantly, the ACTN architecture allows for hierarchy and recursion not only of SDN controllers but also of

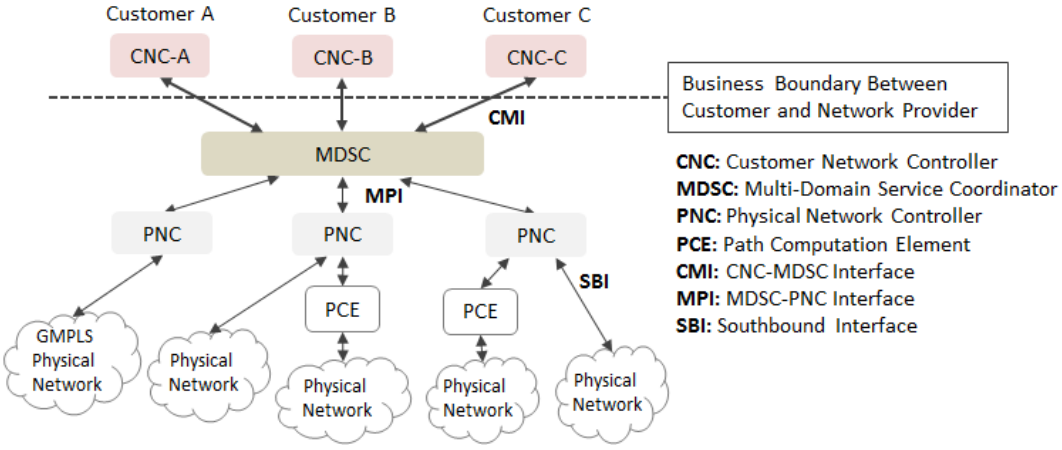

FIGURE 3.2: IETF Abstraction and Control of Transport Networks (ACTN) Architecture.

traditionally controlled domains that use a control plane.

With regard to data models, protocols and APIs, the OpenFlow (OF) protocol standardized by ONF is likely the most popular protocol used in the southbound interface of SDN architectures. The OF specification [42] currently defines two elements: (1) an abstract model of a switch datapath for packet processing (i.e. the expected behavior of a switch); and (2) a protocol for the communication between the switch and the SDN controller to program the behavior of the switch dataplane. While the current scope of OF is basically flow management, the ONF is seeking as future evolutions of the protocol to expand the scope of SDN control, to support a broad spectrum of datapath hardware platforms, including fully programmable packet switches (i.e. switches with no built-in protocol behavior) [43]. Another important initiative within the ONF is the *Information Modelling Project* (ONF-IMP), which intends to provide a common basis for terminology definition and normalization underpinning SDN API development to facilitate convergence of model-based interface definitions. To that end, the ONF-IMP has established the so-called ONF Common Information Model (ONF-CIM) [44], which includes all of the artifacts (objects, attributes and relationships) that are necessary to describe the domain for the applications being developed. The ONF-CIM comprises a core model (ONF Core Information Model [45]), which provides a technology-agnostic representation of network forwarding resources from a management-control perspective, and various specific technology and layer additions (e.g. OTN/OCH/ODU, ETH, MPLS-TP).

The ONF-CIM might be continually expanded and refined over time, to add new applications, capabilities or technologies, or to refine it as new insights are gained. Building on the ONF-CIM, the *Open Transport project* within the ONF addresses the SDN and OF standard-based control capabilities for transport technologies of different types, including optical and wireless transport. The work includes identifying and addressing different use cases, defining the application of SDN architecture and information modeling to transport networks, and defining standard SDN interfaces for transport networks, including OF protocol extensions and transport controller APIs. Three relevant outputs to consider in our discussion from the Open Transport project are: ONF TR-522 [46], which describes the application of the general SDN architecture [40] and the ONF-CIM to transport networks; ONF TR-527 [47], which develops the functional requirements for the definition of a Transport API (T-API); and TR-532 [48], which provides a technology specific extension to the ONF core information model [45] for the use of the SDN architecture in wireless transport networks. Still within the ONF, it's also worth mentioning the *Northbound Interfaces project* that develops concrete requirements, architecture, and working code for northbound interfaces in order to lower barriers to SDN application development. Thus far, only document ONF TR-523 [49] stating the principles for the definition of intent-based interfaces has been produced. Within the IETF domain, YANG [50] is becoming the data modeling language of choice. YANG can be used to model both configuration and operational states. It is vendor-neutral and supports extensible APIs for control and management of elements. Indeed, YANG data models [50], together with appropriate messaging protocol (e.g. NETCONF [51]) or RESTCONF [52]) and encoding mechanisms, have been already adopted and promoted by several industry-wide open management and control initiatives (e.g. OpenConfig). YANG data models are also being considered to provide solutions for the ACTN framework [53]. For more information on SDN architectures and technologies along with key developments within ONF, IETF and other standard development organizations and industrial, the interested reader is referred to [54][55].

3.3 Satellite Network Architecture

A technology-agnostic reference architecture for Broadband Satellite Multimedia (BSM) communications systems has been established by ETSI [56]. The BSM system architecture is conceived as an overarching architecture consisting of the common components found in an interactive satellite communications network: User Satellite Terminal (ST), Gateway ST, satellite payload, Network Management Centre (NMC) and Network Control Centre (NCC). Importantly, the BSM system architecture is not restricted to any particular satellite air interface (e.g. DVB-S2/RCS2) but intended to support diverse air interface protocols. Indeed, the overall ETSI BSM system architecture is applicable to the different configurations that a satellite network can be implementing in terms of topology (star, mesh) and payload operation (transparent and regenerative) [57].

Figure 3.3 depicts the ETSI BSM architecture in terms of reference interfaces for the user plane (U-plane) and for the control / management planes (C-plane and M-plane). The reference interfaces are divided into physical and logical interfaces, the former referring to physical connections between equipment and the latter referring to logical associations between peer protocol entities. As illustrated in Figure 3.3, one central principle of the BSM system architecture is the logical separation of the Satellite Independent (SI) layers (e.g. Ethernet/IP layers together with the interworking and adaptation functions needed for the interconnection with external networks) from the Satellite Dependent (SD) layers, whose interaction is formalized by the definition of a Satellite Independent - Service Access Point (SI-SAP) interface [58]. Focusing on the U-plane (aka data plane), 4 physical interfaces are identified at the interconnection points between the premises network and the User ST (T interface), User ST and satellite payload (U/U_{ST} interface), satellite payload and Gateway ST (U/U_{GW} interface) and Gateway ST and external network (G interface). The radio interface label U means that the User ST and Gateway ST have the same radio interface to communicate among them through the satellite payload while U_{ST} and U_{GW} refers to the case that the radio interface is different in the two sides. On the other hand, three logical interfaces are defined for the U-plane, corresponding to the peer-to-peer interactions of the different layers of the radio interface protocols.

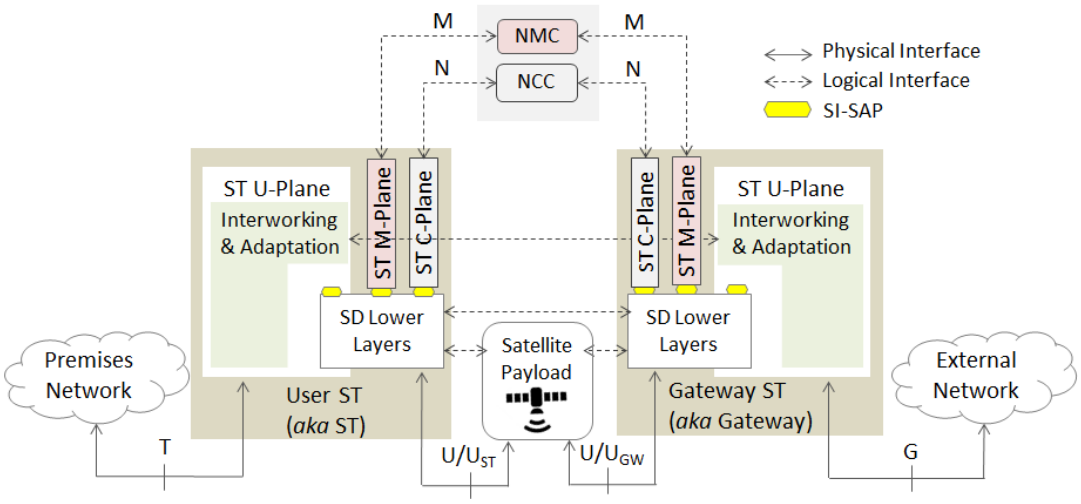

FIGURE 3.3: ETSI BSM System Architecture: Reference Interfaces for U/C/M-Planes.

One logical interface covers the interaction between SI protocol layers at both sides, i.e. the interworking and adaptation functions. The other two logical interfaces fit within the SD lower layers, one for interfacing with the satellite payload and another for the peer ST. The boundary between these two logical interfaces depends on the supported satellite payload capabilities. With regard to the C-plane and M-plane, two logical interfaces named *N* and *M* are identified. In particular, interface *N* is a control interface between the User/Gateway STs and the NCC, which is the functional entity that provides the real time control of the BSM network (e.g. session/connection control, routing, terminals access control to satellite resources, etc.). And interface *M* is a management interface between the STs and the NMC, which is the functional entity in charge of the management of all the system elements in the BSM network (e.g. Fault, Configuration, Performance, Accounting and Security [FCAPS] management). Of note is that currently both *N* and *M* interfaces are considered as internal interfaces within the BSM system, not subject to standardization or harmonization between vendors. However, we devise this functional separation established in the BSM reference model as the foundational point to introduce SDN concepts and technologies within the BSM system, as detailed later on in this chapter.

With regard to the BSM service capabilities and QoS support over the satellite links,the BSM system architecture defines BSM bearer services. A *BSM bearer service*

includes all aspects to enable the provision of a U-plane data transport service between the User/Gateway STs, including the QoS characteristics and other properties such as connectionless or connection-oriented, unidirectional or bidirectional, symmetric or asymmetric, point-to-point / multicast / broadcast nature of the bearer service. The BSM bearer services are defined at SI-SAP interface level and use the services provided by the underlying native bearer services (which depends on the specific implementation of the SD lower layers for link and medium access control). In the same way, the higher layer services (e.g. IP connectivity over the satellite network) are built on the BSM bearer services and can be mapped to different BSM bearer services depending on the particular higher layer service requirements. The abstract representation of the available BSM bearer services at SI-SAP level is done via labels called Queue Identifiers (QIDs). The QoS properties associated with a given QID are defined by QoS specific parameters and each QID is mapped onto suitable lower layer transfer capabilities in order to realize that QoS. QIDs are defined in more detail in the SI-SAP specification [58] and SI-SAP Guidelines [59]. The QoS model established for BSM systems and the traffic classes used to describe QoS, performance management and resource allocation are defined in detail in [60][61], respectively.

3.4 SDN-enabled Satellite Network Architecture

Based on the previously described aspects of the BSM system architecture (i.e. functional components, reference interfaces, bearer services/QIDs, QoS model), Figure 3.4 illustrates the defined solution for the adoption of an SDN architecture within the satellite network [62].

This solution relies on the introduction of an SDN controller as part of the satellite network functional architecture to manage the connectivity services between the *T* and *G* reference points. In the case of packet switched services (e.g. IP and Ethernet connectivity services), the finest granularity for QoS forwarding treatment is commonly referred to as a *flow*, which can be defined as a sequence of packets between a source and a destination intended to receive identical service policies when progressing through the U-plane. A set of packet filters, referred to as Traffic Flow

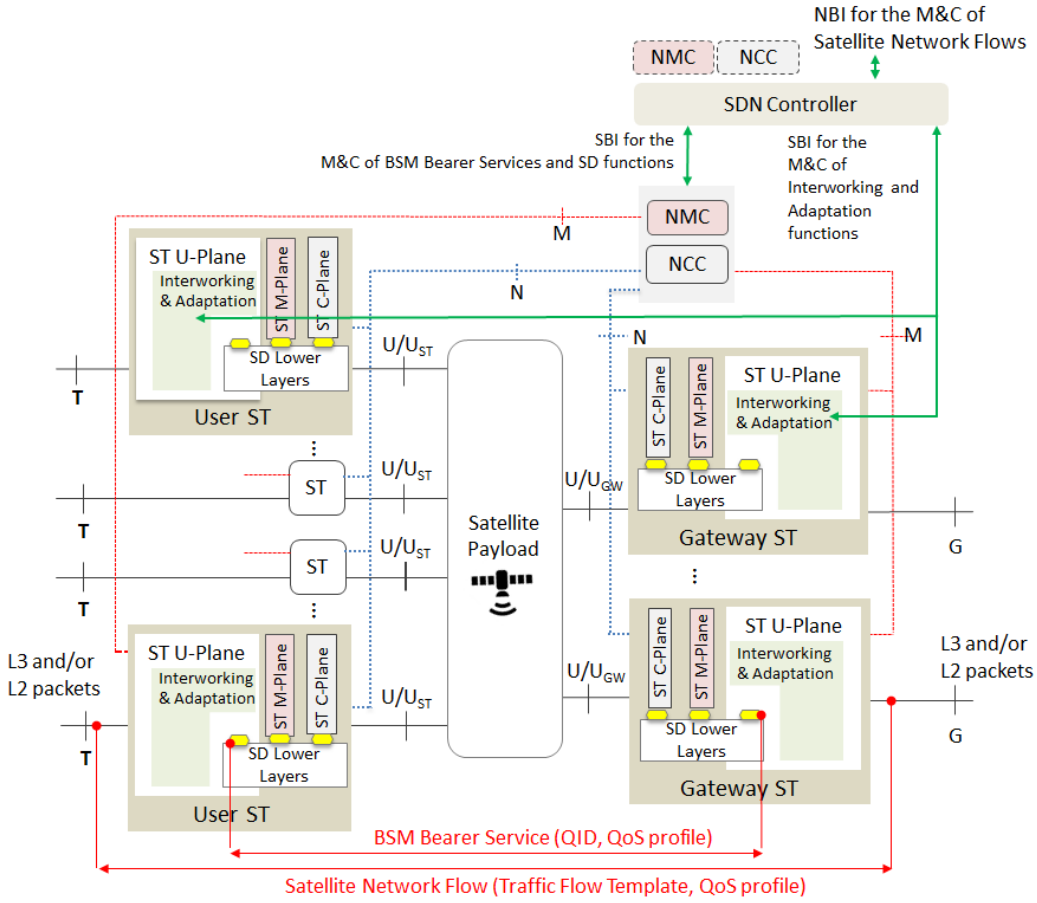

FIGURE 3.4: SDN-based Satellite Network Architecture procedures (Source: VITAL deliverable D4.5 [62]).

Template in Figure 3.4, shall be used to identify individual data flow belonging to a specific application (e.g. the packet filters for IP flows typically consist of IP 5-tuples with source IP address, destination IP address, source port, destination port and protocol type). As depicted in Figure 3.4, the SDN controller directly manages the satellite independent services such as the IP/Ethernet layer QoS and indirectly manages the SD services through the NCC/NMC functions. Accordingly, the following interfaces are then needed:

- SBI for the management and control (M&C) of the interworking and adaptation functions in the Gateway STs and potentially also in User STs. This interface is not satellite dependent so that SDN models and interfaces used in the broad networking domain can be adopted such as *OF* and *YANG models*.

- SBI for the M&C of the BSM bearer services and potentially also of some capabilities within the SD lower layers (satellite resources such as a frequency plan, modulation and coding schemes or other satellite specific properties) through the interaction with legacy satellite network NCC/NMC functions. This interface may have to consider satellite specific aspects so that some extension and adaption of existing SDN models and interfaces is necessary. Potential candidate baseline SDN data models and interfaces for the realization of this interface are *OF* and the *Microwave Information Model* [48]. In case NCC/NMC functions could be eventually implemented as network applications on top of the SDN controller, another potential solution for this interface could be based on an extension of the *ETSI SI-SAP interface* for the realization of the *N* and *M* interfaces directly serving as SBIs from the SDN controller viewpoint.

- NBI for the management and control of the satellite network flows by network applications running on top of the SDN controller or from external controllers within an upper-level control domain. Potential candidate SDN data models and interfaces for the realization of this interface are *OF*, the *ONF Transport API* [45] and the *YANG models* as identified for the ACTN architecture.

3.5 SDN Data Models and Interfaces

The main characteristics and pros/cons of the previously mentioned candidate data models and interfaces for consideration within the SDN-based satellite network architecture are discussed in the following:

ETSI BSM SI-SAP

The SI-SAP interface provides a functional separation between satellite dependent (SD) and independent (SI) layers. The SI-SAP interface is currently specified in terms of the primitives exchanged between the SD and SI layers, following ISO/OSI protocol stack model. The existing specification [63] defines primitives to support U-plane and C-plane functionalities. More specifically, the C-plane services provided by the SI-SAP interface are: (1) *Logon/logoff services;* (2) *SI layer configuration service*, to provide the SI layer with the necessary information to configure e.g. the addressing plan and different functions of higher protocol layers such as header compression; (3) *Address resolution service*, used to perform the address mapping between SI and SD layers; (4) *Resource reservation service*, used for resource allocation and overall QoS management; and (5) *Multicast group receive and transmit services*, invoked to build multicast groups and to receive the desired multicast data flows.

The SI-SAP could be deployed as an external interface, as considered in [63]. Under such an approach, the SD lower layers and the control entity can be running at different places and be interconnected by point-to-point protocol technology (e.g. Ethernet). As such, the SI-SAP interface service primitives are defined as specific messages transported by the technology implemented by the point-to-point protocol. Message formats and protocol encapsulation options are discussed in [63]. Additional information on the use of the SI-SAP interface can be found in [59].

Therefore, the BSM SI-SAP is a clear candidate for the implementation of SBIs for the management and control of the BSM bearer services and potentially also some capabilities within the SD lower layers. However, to that end, current specifications should be revisited and extended since they are not currently conceived to manage physical radio aspects of the satellite link (e.g. modulation and coding scheme selection) and have limited monitoring and management plane capabilities.

ONF OpenFlow

OF fundamentally provides a solution for flow management, allowing a fine-grained control of the forwarding behavior at packet level of a switch/router node. The OF specification [42] currently defines two elements: (1) an abstract model of a switch datapath for packet processing (i.e. the expected behavior of a switch); and (2) a protocol for the communication between the switch and a controller to enable the controller to program the behavior of the switch dataplane. The OF specification features support for a number of commonly used dataplane protocols ranging from Layer 2 to Layer 4, with packet classification being performed using stateless match tables and packet processing operations (called actions or instructions) ranging from header modification, metering, QoS, packet replication (e.g. to implement multicast or link aggregation) and packet encapsulation / decapsulation. The specification also counts with several artefacts for statistics collection, which can be retrieved on demand or via notifications. A complementary protocol to the OF protocol is OF-CONFIG, also standardized by ONF. OF-CONFIG adds configuration and management support to OF switches. OF-CONFIG provides configuration of various switch parameters that are not handled by the OF protocol.

However, current OF specification has very limited support to cope with physical layer aspects of the switch ports. Thus far, the OF specification has only introduced a set of port properties to add support for optical ports. As to the support of wireless ports, the only consideration has been to define the process for sending a packet through the same port that it is was received, a behavior that was not clearly defined in in earlier versions of the specification and that is typical in wireless links. Therefore, there is no practical support in current OF specification for the configuration and monitoring of wireless links/ports.

Accordingly, within the SDN-based satellite network, OF is a clear candidate to be used internally to control the switching functions within Gateways and STs. Remarkably, OF could also be used as an external interface to expose some control for satellite network flow management and so provide the control features necessary for the realization of E2E TE (this approach is the one further elaborated in chapter 6). Of note is that the exposure of an OF interface by a satellite network has been also proposed in the context of the realization of Virtual Network Operator (VNO)

solutions [64], in which a VNO is provided with an interface to control and manage the satellite segment resources leased from a satellite network operator as if it was programming an OF switch. All in all, OF is an extensible protocol, providing mechanisms for SDN programmers to define additional protocol elements (e.g., new match fields, actions, port properties, etc.) to address new network technologies and behaviors (i.e., the protocol defines the expected behavior of the switch but also how the behavior can be customized using the interface).

ONF Microwave Information Model

The Microwave Information Model (IM) is an effort led by the ONF to define a standard of a common and generic information model for SDN-enabled wireless transport environments in order to simplify the operations and control of microwave/millimetre wave radio link network elements and facilitate the integration of distinct multi-vendor solutions under a common and single control framework. The Microwave IM is provided in ONF TR-532 [48] as a technology-specific extension to the TR-512 ONF CIM that can be implemented as a YANG data model so that the control-management of the microwave device by the SDN controller can be realized via the NETCONF protocol. The Microwave IM provides the necessary attributes for the device informing the SDN controller about its capabilities, the controller configuring the device, and the device providing status, problem and performance information. For example, the Microwave IM allows for the configuration of frequency plans (channel arrangement), channel frequencies and transmission bandwidths, used modulation schemes, etc.

The current specification is limited to point-to-point radio-links. However, this model could be a valid starting point to base a model for the satellite physical layer and be used as an internal SBI for the management and control of the SD lower layers (satellite resources such as a frequency plan, modulation and coding schemes or other satellite specific properties).

ONF Transport API

The ONF Transport APIs (T-APIs) seeks to provide programmable access to transport SDN controller functions by abstracting a common set of control plane functions such as network topology, connectivity requests and path computation to a

set of service interfaces. T-APIs are intended to be applicable on the interface between a Transport SDN controller "Black Box" and its client application. The actors involved in the information exchange over this interface include transport network provider domain controllers in the role of producers and the transport network application systems in the role of the consumers. The transport network application systems could be either a business client system (which itself may include some control functions) or the network operator's upper level control, orchestration and/or operations systems. The T-APIs are also intended to be equally applicable between the controllers within a transport controller recursive hierarchy. The expected services delivered by the T-APIs are:

- *Topology service:* API to retrieve network topology, node, link, and node edge points.
- *Connectivity service:* API to request create, update, and delete connectivity including point-to-point and multipoint.
- *Path computation service:* API to request computation and optimization of paths.
- *Virtual network service:* API to create, update, and delete virtual network topologies.
- *Notification service:* API to support publish/subscribe models for asynchronous notification of events such as failures or degradations.

While ONF T-APIs development is still work in progress (thus far, ONF document TR-527 [47] only provides the functional requirements for the specification of T-APIs), it could be a clear candidate for the realization of a NBI for the M&C of the end-to-end flow service delivered by a satellite network.

YANG Models

YANG models have been produced to allow configuration or modeling of a variety of network devices, protocol instances, and network services. A classification of YANG data models is given in [65], the latter reference more focused on service models. In particular, four types of service YANG models are distinguished:

- *Customer Service Model:* A customer service model is used to describe a service as offer or delivered to a customer by a network operator.

- *Service Delivery Model:* A service delivery model is used by a network operator to define and configure how a service is provided by the network.
- *Network Configuration Model:* A network configuration model is used by a network orchestrator to provide network-level configuration model to a controller.
- *Device Configuration Model:* A device configuration model is used by a controller to configure physical network elements.

YANG models coupled with the RESTCONF/NETCONF protocol provides solutions for the ACTN framework, which indeed seeks to provide a control hierarchy and interfaces that would enable deployment of multi-domain transport SDN networks. Hence, according to [66], *Customer Service Models* would be applicable to the ACTN CMI interface, *Network Configuration Models* to the MPI and *Device Configuration Models* to the SBI. In this context, and considering that the integration of the proposed SDN controller of the satellite network within an ACTN architecture would likely be realized through an MPI interface, existing YANG models applicable in the MPI interface that are not OTN/WSON technology specific are summarized in Table 3.1. Note that various YANG models are work in progress. Furthermore, there is also IETF Internet Draft [67] aimed to describe use cases that could be used for analyzing the applicability of the existing models defined by the IETF for transport networks with a focus on MPI interface.

3.6 Integration Approach for E2E SDN-Based TE in Satellite-Terrestrial Backhaul Networks

The exposition of control and management capabilities of the satellite connectivity services through an SDN-based interface would allow a Mobile Network Operator (MNO) to easily integrate and operate the satellite component within a backhauling infrastructure that is progressively relying on SDN technologies for the terrestrial capacity counterpart. Managing both terrestrial capacity and satellite capacity under a centralized and consistently operated SDN framework enables the deployment of end-to-end SDN-based traffic engineering (TE) solutions.

TABLE 3.1: YANG Models for Traffic Engineering.

Function	**Yang Model**
Configuration Scheduling	X. Liu, et. al., "A YANG Data Model for Configuration Scheduling", draft-liu-netmod-yang-schedule, work in progress.
Path computation	I.Busi, S.Belotti et al. "Path Computation API", draft-busibel-ccamp-path-computation-api-00.txt, work in progress
Path Provisioning	T. Saad (Editor), "A YANG Data Model for Traffic Engineering Tunnels and Interfaces", draft-ietf-teas-yangte, work in progress.
Topology Abstraction	X. Liu, et. al., "YANG Data Model for TE Topologies", draft-ietf-teas-yang-te-topo, work in progress.
Tunnel PM Telemetry	Y. Lee, D. Dhody, S. Karunanithi, R. Vilalta, D. King, and D. Ceccarelli, "YANG models for ACTN TE Performance Monitoring Telemetry and Network Autonomics", draft-lee-teas-actn-pm-telemetry-autonomics, work in progress.
Service Provisioning	No references available yet

TE mechanisms are used to optimize the performance of a data network by dynamically analyzing, predicting, and regulating the behavior of the traffic across the network [68]. In integrated satellite-terrestrial backhaul network, TE solutions shall be able to use the satellite capacity in the way that best complements the terrestrial capacity in front of the changing conditions of both traffic demand (e.g. increase of traffic demand for an especial event, spatial demand fluctuations over time) and network situation (e.g. backhaul backup for terrestrial link failures, network rapid roll-out, fast response capacity, cells on wheels). Facing these multiple and diverse conditions in a consistent manner becomes challenging for traffic engineering. Compared to the traditional MPLS/TE mechanisms used in today's transport networks, the big advantage of a centralized SDN framework for the realization of TE solutions is that there is a holistic view of the network together with mechanisms to enforce network polices from a single touch point [69].

A network architecture framework for the realization of E2E SDN-based TE in satellite-terrestrial backhaul networks is presented in the following, together with a couple of illustrative TE workflows to validate the proposed integration approach.

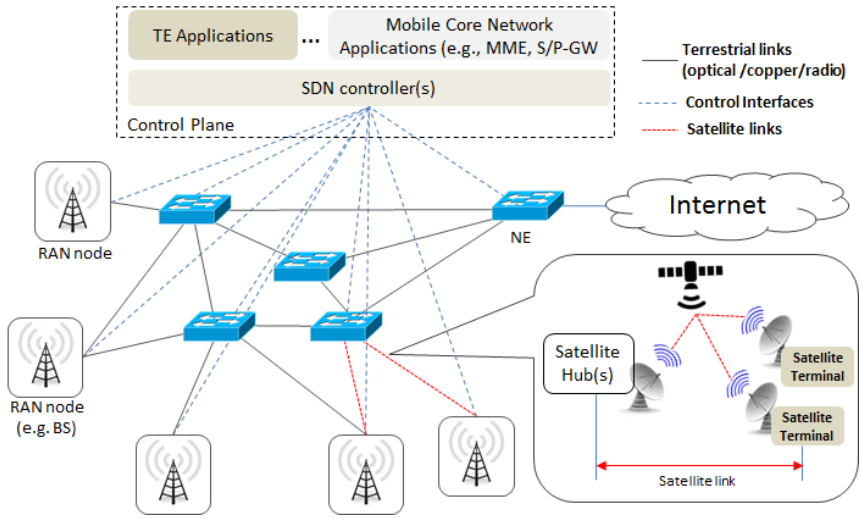

FIGURE 3.5: SDN-Based Hybrid Satellite-Terrestrial Mobile Backhaul Network.

3.6.1 Network Architecture Framework for Hybrid Satellite-Terrestrial Mobile Backhaul Network

Several proposals exist for adopting SDN concepts in mobile network architectures [70][71]. In general terms, an illustrative view of an SDN-based mobile that uses SDN-enabled transport from the RAN nodes (e.g. BSs) all the way through the backhaul to the core network is depicted in Figure 3.5. Though this architecture is contextualized for LTE technology, this vision is claimed to be generic and not constrained by the specifics of the LTE standard.

As depicted in Figure 3.5, Mobile Core Network (MCN) control functions (e.g. Mobility Management Entity [MME] and Serving / Packet Data Network [PDN] gateways (S/P –GW) functional elements in LTE Evolved Packet Core [EPC]) together with specific TE functions for the transport network are realized as applications running on top of an SDN controller (represented here as a single functional entity but likely to follow a hierarchical structure of controllers). This SDN controller is responsible for managing the Network Elements (NE) that provide the packet switching and forwarding capabilities within the transport network. In this respect, the underlying transport network infrastructure may involve a number of different physical network equipment, or forwarding devices such as routers, switches, virtual switches, to name a few. Building on the above view of SDN-based mobile

networks and on the SDN-based satellite network architecture discussed in previous section, Figure3.6 depicts the functional view of the proposed integration approach, which is founded on two main concepts:

- *Abstraction of the overall satellite network as an SDN-capable "switch"*. In particular, the OF switch abstraction model [64] is considered to model the operation of the satellite network as seen from the MNO SDN controller entity. This corresponds to one of the candidate solutions discussed in the previous section for the realization of the NBI interface for the control and management of the satellite network connectivity services.

- *Use of SDN-based TE applications*, with a central Path Computation Engine (PCE) that support the operation of the MCN applications for traffic management within the backhaul transport network. It is assumed that the overall transport network is managed as a single logical forwarding domain and that, inside the forwarding domain, a MNO's SDN controller makes the forwarding decisions. As depicted in Figure 3.6, all SDN-capable L2/L3 NEs are connected to the MNO's SDN network controller through OF interfaces, including the "Satellite network switch". In this way, SDN-based TE mechanisms can seamlessly span the whole network. For the terrestrial connection, no specific technology is assumed rather than considering that traffic flows can also be managed through SDN features.

In order to raise different considerations with regard to the operation of TE procedures, the illustrative network topology depicted in Figure 3.6 considers three RAN nodes with LTE eNB functions, one connected to the transport network only by terrestrial means (RAN node#C), another connected only through the satellite network (RAN node#A) and a third one (RAN node#B) connected to both a terrestrial connection and a satellite connection through an SDN-capable Cell Switch Router (CSR). This third case is used to illustrate the realization of TE mechanisms for multi-path optimization. With respect to the terrestrial part of the transport network, three NE are included in the reference network topology, two of them acting as internal aggregation/core nodes within the transport network (i.e. NE#A and NE#B) and the third one (i.e. NE#C) providing the interconnection with the external networks (e.g.

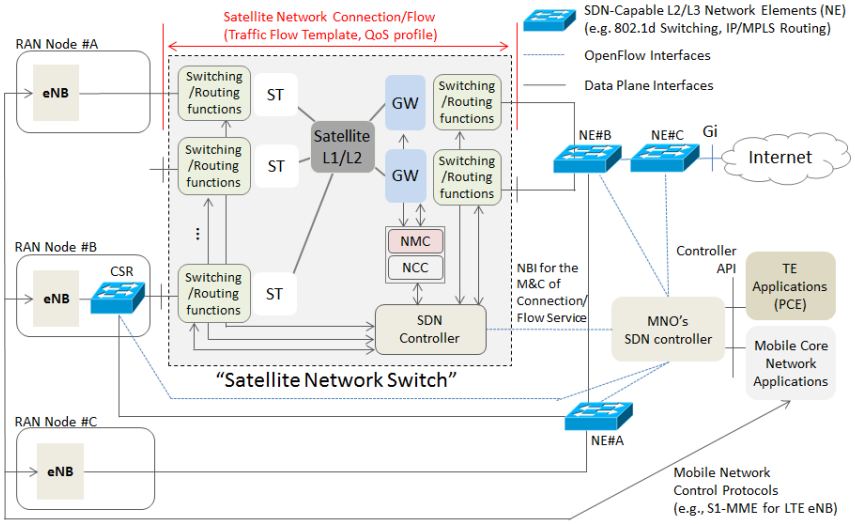

FIGURE 3.6: Functional View and Illustrative Network Topology Used in the TE Workflows (Source: VITAL deliverable D4.2 [72]).

Internet) through a conventional 3GPP Gi interface. Of note is that, in addition to OF interfaces for controlling the forwarding function of the transport network, other control interfaces are likely to be in place in the overall setting for other purposes, such as the 3GPP S1-MME interface between the MCN applications and the eNBs within the RAN nodes to manage the activation/deactivation of radio access bearers in the eNB for the served mobile terminals.

3.6.2 Illustrative TE Workflows

Two illustrative workflows are presented to validate the proposed integration approach. The first one shows the activation of a traffic flow through the satellite-terrestrial network to enforce a mobile network bearer (e.g. so-called Evolved Packet System [EPS] bearer in the context of LTE) that can benefit from optimal path computation. The second workflow shows the modification of an already established flow as a reaction to a congestion/failure situation in one link within the transport network.

Flow activation with optimal path computation

Based on the network topology depicted in Figure 3.6, a message chart with the operation of a path computation mechanism for multi-path satellite-terrestrial traffic optimization is provided in Figure 3.7.

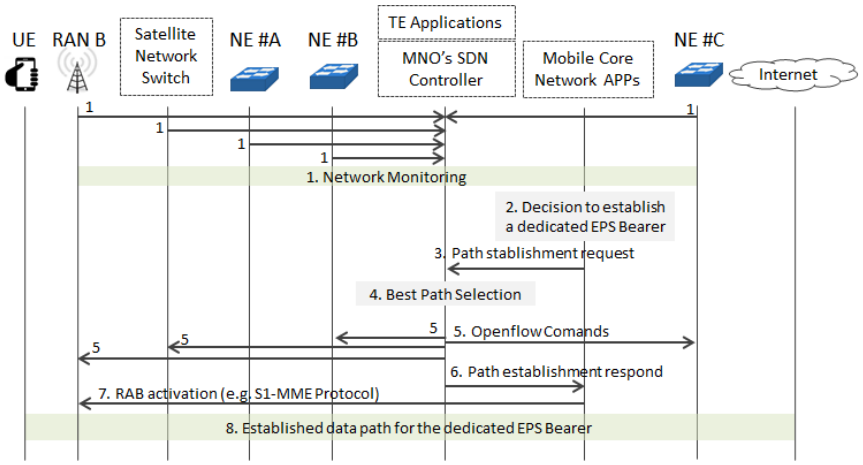

FIGURE 3.7: Flow Activation with Optimal Path Computation.

In particular, the provided workflow covers the case of the establishment of a dedicated EPS bearer that relies on the TE path computation mechanism to activate the traffic path between the RAN node and the external network reachable through NE#C considering the characteristics of the EPS bearer and the load conditions across the whole network. It is assumed that the SDN controller has a global view of network topology which can be represented by means of a graph including all links between OF switches (the links could be found by e.g. leveraging protocols such as LLDP (802.1AB), which is used by network devices to advertise their identity, capabilities, and neighbours). Details of the different steps depicted in Figure 3.7 are given in the following:

- ***Step 1:*** Monitoring of the SDN forwarding elements within the domain, including the CSR, "Satellite Network Switch (SNS)" and NEs. Solutions such as the one described in [73] allows for an OF controller to have accurate monitoring of per-flow throughput, packet loss and delay metrics in order to aid TE. In this respect, while a flow is active, the controller and the SDN forwarding element can exchange messages concerning the state of the flow.

- ***Step 2:*** As a result of the activation of a new service (e.g., HD video-streaming service) by a mobile terminal connected in RAN node#B, the MCN decides to establish a new dedicated EPS bearer to support that service. The activation of the dedicated EPS bearer requires the activation of a flow with QoS guarantees across the transport network. The two edge nodes of the EPS bearer are

the RAN node#B, where the UE is assumed connected, and the NE#C, which serves as the gateway to the external network.

- ***Step 3:*** The MCN request to the TE application the computation of the best path between RAN node#B and NE#C. QoS attributes of the EPS bearer are indicated (e.g. Guaranteed Bit Rate).

- ***Step 4:*** Based on the (1) network topology knowledge, (2) the network monitoring information and (3) QoS attributes of flow, the TE application can compute the most appropriate path. Different algorithms could be supported here, including graph searching algorithms for path finding and algorithms for path selection depending on policies with respective of traffic engineering or service quality, such as calculating the shortest path forwarding based on a consistent view of network state or provision application-aware routing [74]. Anyway, let us consider that the outcome of this decision is that a path through the satellite network is chosen for this flow.

- ***Step 5:*** Flow entries are installed into the OF switches by the MNO's SDN controller so that traffic associated with the EPS bearer is forwarded through the selected path.

- ***Step 6:*** The MCN gets the path establishment response.

- ***Step 7:*** The EPS bearer activation at the radio layer takes places (i.e. Radio Access Bearer [RAB] activation), involving the interaction between the MCN functions and the eNB within RAN node#B.

- ***Step 8:*** The data plane for the dedicated EPS bearer gets live and traffic follows the selected path through the satellite network.

The above workflow assumes that the path is established to support a single EPS bearer. However, the same approach would be used in case of deciding the best path for traffic aggregates with common QoS requirements. This is well supported in OF by just establishing the corresponding matching conditions (e.g. IP prefixes to identify a traffic aggregate in front of particular IP addresses of the individual flows).

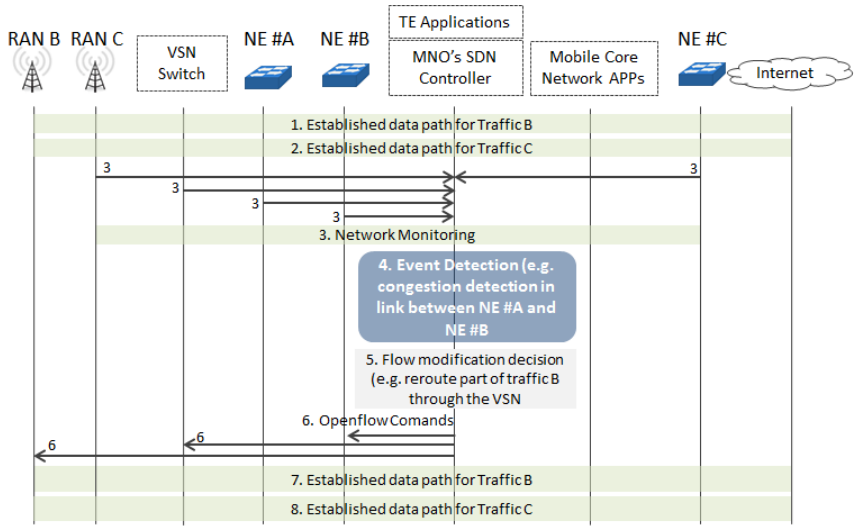

FIGURE 3.8: Flow Update to Overcome Congestion/Failures.

Flow update to overcome congestion/failures

The Figure 3.8 shows how the failure of a path, or simply the congestion of a path that could cause QoS degradation, could be handled within the proposed integration approach.

In particular, the message chart depicted in Figure 3.8 is a TE mechanism that will update an already established flow in order to overcome a congestion/failure event. Details of the different steps depicted in Figure 3.8 are given in the following:

- ***Step 1:*** The starting point considers that traffic from/to RAN node#B and from/to RAN node#C, called Traffic B and Traffic C respectively, are both flowing through NE#A, NE#B and NE#C. This could be assumed to be the optimal traffic path for a moderate traffic load scenario.

- ***Step 2:*** Monitoring of the SDN forwarding elements is conducted by the MNO's SDN controller, as described in the previous workflow.

- ***Step 3:*** An event that puts at risk the QoS of the established flows occurs. This could be, for example, a considerable traffic increase in RAN nodeC at certain time of the day that overloads the link among NE#A and NE#B, which is shared by Traffic B and Traffic C.

- ***Step 4:*** The TE application detects the congestion situation. For example, the TE application could have set a high utilization threshold of 60 percent and

low utilization threshold of 20 percent for the traffic load on the shared link. If this high threshold is exceeded, high utilization is observed and e.g. a part of Traffic B could be switched to pass through the satellite network.

- ***Step 5:*** Flow entries are installed to OF switches along the path by the MNO's SDN controller to re-route part of the traffic B through the satellite connection.
- ***Step 6:*** While the path for traffic C remains unchanged, now part of traffic B is served through the satellite network, reducing congestion in the link between NE#A and NE#B.

Flow updates can also be driven by connection protection in case of failure. Indeed, path protection and network recovery from failure are critical aspects of TE. While these aspects are well-understood in conventional MPLS/IP networks, work is still needed to mature these concepts in the context of SDN networks [75].

3.6.3 Summary

This chapter has elaborated on the support of SDN concepts and technologies within satellite networks and developed a case study for the applicability of SDN-based TE solutions for the management of a satellite component integrated in next-generation mobile backhaul networks. Building upon the ETSI functional architecture for BSM systems, it has been explained a solution for the adoption of an SDN architecture within the satellite network. On this basis, an integration approach for the realization of E2E SDN-based TE in satellite-terrestrial backhaul networks has been presented in which the satellite component has been abstracted as an OF switch. Two central TE workflows have been developed to validate the proposed integration approach.

Based on the network architectures definitions discussed in this chapter, the following chapters will present the capacity analysis and the traffic distribution strategies developments for hybrid satellite-terrestrial mobile backhaul networks.

Chapter 4

Backhaul Capacity Gains of Traffic Overflow Through Satellite

4.1 Introduction

Mobile backhaul networks are characterized by a large number of small/medium sites connected to centralized aggregation points in charge of concentrating traffic towards the MNO core network. The centralized control and programmability capability of SDN deployed across all the different technologies and networking layers co-existing in mobile backhaul networks would greatly simplify its operation and make network resource management more flexible and efficient (e.g. capacity-aware end-to-end path computation). However, while capacity and traffic management aspects (e.g., capacity dimensioning, routing, QoS, congestion, resilience) have been extensively analysed for mobile terrestrial backhaul networks [23]-[26], little attention has been paid to these aspects when it comes to the consideration of hybrid terrestrial-satellite networks for mobile backhauling.

In this context, in this chapter is proposed a model for the analysis of capacity and traffic management strategies in hybrid satellite-terrestrial mobile backhauling networks that rely on SDN for fine-grained traffic steering. Numerical results are provided to assess the capacity gains when the satellite backhaul capacity is used for traffic overflow, taking into account the placement of the satellite capacity at different traffic aggregation levels and the spatial correlation of the traffic demand. The rest of the chapter describes the system model and problem formulation to cope with the analysis of different capacities and traffic management strategies of the satellite

component and the assessment results are presented.

4.2 Network Model

Let's consider a cellular network with M Base Stations (BS) deployed across a large geographical area. A terrestrial backhaul network is assumed to provide the primary connectivity between the m-th base station (BS_m) and the mobile core network, likely combining diverse transmission technologies over fiber, copper and/or microwave links. This terrestrial backhaul network is considered to be structured in multiple levels of traffic aggregation. The N traffic aggregation points at the first aggregation level are referred to as $A1_n$, $n \in \{1, ..., N\}$, each one anchoring the traffic of a number of geographically nearby BSs. Complementing the terrestrial backhaul network, a total of K satellite terminals (ST) denoted as ST_k, k$\in \{1, ..., K\}$, are assumed to be deployed at some of the BS sites and/or $A1$ locations to provide supplemental connectivity through a satellite backhaul network. It is worth noting that satellite backhaul connectivity at higher traffic aggregation points above $A1$ is assumed to be a non-feasible option due to the high traffic volume that would likely be necessary to convey through the satellite link at specific locations. Therefore, it is considered that a given ST_k is used either to backhaul traffic from a single BS or to serve the aggregated traffic of multiple nearby BSs connected to the same $A1$ aggregation point. An illustrative view of the network model is depicted in Figure 4.1.

The terrestrial backhaul network topology between the BS and $A1$ aggregation points is captured with a terrestrial connectivity vector $\pi = \{\pi_1, ..., \pi_M\}$, where $m = n \in \{1, .., N\}$ indicates that BS_m is attached to the $A1_n$. The case of a deployable BS not reachable from terrestrial backhauling connectivity is captured by setting the corresponding coefficient to zero. Similarly, the satellite backhaul network topology is captured with a satellite connectivity vector $\Omega = \{\Omega_1, ..., \Omega_{p=M+N}\}$, where $p = k \in \{1, ..., K\}$ for $p \leq M$ indicates that BS_p counts with satellite connectivity through the ST_k, and $p = k \in \{1, ..., K\}$ for $M < p \leq N$ indicates that $A1_{p-M}$ has satellite connectivity through the ST_k.

Capacity limitations are assumed to be potentially found at; (1) terrestrial backhaul connections between BS and $A1$ locations; (2) terrestrial backhaul connections

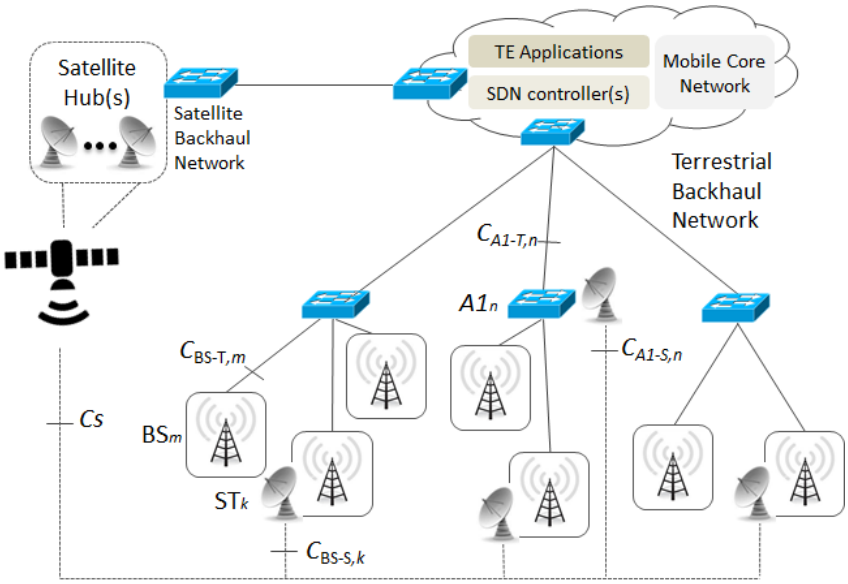

FIGURE 4.1: System Model for the Hybrid Satellite-Terrestrial Backhaul Network.

between $A1$ to higher aggregation points and; (3) satellite backhaul connections at each ST. Therefore, three capacity vectors are defined to represent the maximum capacity available at each of these connectivity points to the backhaul network: $\mathbf{c}_{BS-T} = \{C_{BS-T,1}, ..., C_{BS-T,M}\}$, $\mathbf{c}_{BS-S} = \{C_{BS-S,1}, ..., C_{BS-S,K}\}$, $\mathbf{c}_{A1-T} = \{C_{A1-T,1}, ..., C_{A1-T,N}\}$ and $\mathbf{c}_{A1-S} = \{C_{A1-S,1}, ..., C_{A1-S,N}\}$ that correspond to the maximum terrestrial and satellite capacity in each BS as well as the maximum terrestrial and satellite capacity in each $A1$ aggregation point respectively. In addition, in the case of the satellite connections, it is assumed that the K STs are sharing a common pool of satellite capacity denoted as C_S, so that the actual aggregate capacity in use at a given time through all of the STs cannot exceed C_S.

4.3 Generation of Spatially Correlated Traffic Demand

The traffic demand is captured by computing the vector $\mathbf{t} = \{T_1, .., T_M\}$. This computation needs to account for the degree of correlation that might exist across the components of the traffic demand vector because traffic distribution in mobile networks is shown to exhibit spatial-correlated traffic patterns [76][77].

Accordingly, building upon the traffic model provided in [78][79], the spatial correlation is captured by dividing the coverage area of the mobile network into

AxB parts, forming a grid $g_{a,b}$ in which each part represents the area covered by a single BS. Then, denoting the two-dimensional cartesian coordinates of the center of the grid $g_{a,b}$ as $x_{a,b}$ and $y_{a,b}$, a Gaussian random field $\rho^G_{a,b}$ with $a = 1, ..., A$ and $b = 1, ..., B$, is generated by:

$$\rho^G_{a,b} = \frac{1}{L} \sum_{l=1}^{L} \cos\left(h_l \cdot x_{a,b} + \varphi_l\right) \cos\left(j_l \cdot y_{a,b} + \psi_l\right) \tag{4.1}$$

where the angular frequencies h_l and j_l are uniform random variables between 0 and ω_{max} and phases φ_l and ψ_l are uniform random variables between 0 and 2π. For a large enough L (e.g. 10), the $\rho^G_{a,b}$ variables follow a standard Gaussian distribution with a degree of spatial correlation that can be adjusted through the maximum angular frequency ω_{max}. On this basis, and considering that traffic demand in each BS can be modeled with a log-normal distribution [78][79], the traffic demand T_m for the BS_m at location $g_{a,b}$ is obtained as follows:

$$T_m = exp\left(\sigma \rho^G_{a,b} + \mu\right) \tag{4.2}$$

being μ and σ, respectively, the mean and standard deviation of the natural logarithm of the log-normal distribution. Accordingly, the mean of T_m can be expressed as:

$$\overline{T} = E[T_m] = exp\left(\sigma^2/2 + \mu\right) \tag{4.3}$$

In our analysis, to determine the values of μ and σ, first we fix the ratio (σ/μ), to account for the extent of dispersion of the traffic demand and then we fix the mean $\overline{T}$ so that the traffic volume can be scaled to the values of interest for the analysis. In particular, values in the range of 2% and 20% are considered for the ratio (σ/μ), in line with values provided in the literature [79], and $\overline{T}$ is set to 100 Mbps per BS, in line with typical mobile operator average throughput for 3 sector sites [80]. On this basis, Figure 4.2 illustrates the probability density function (PDF) of the generated traffic demand per BS.

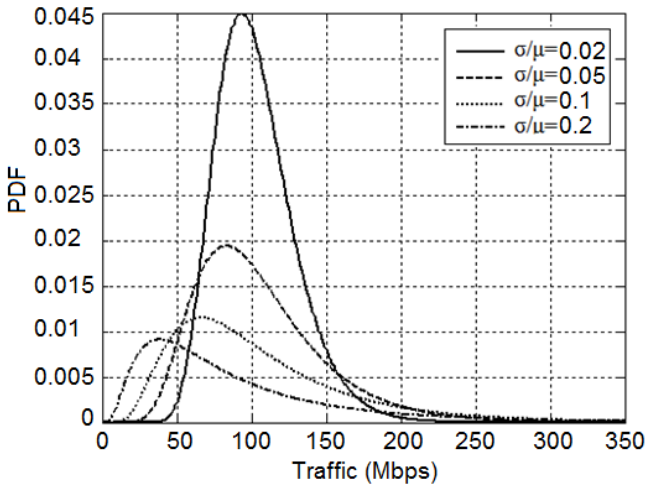

FIGURE 4.2: Traffic Demand for Different Ratios σ/μ and $\overline{T} = 100$ Mbps.

We define the spatial correlation factor as the ratio between the Coherence Distance (CD), which is computed as the distance where the two-dimensional autocorrelation function (ACF) of the traffic demand drops to the half of its peak value, and the Intersite Distance (ISD), which is a common measure to define the density of a cellular deployment. Three different spatial correlation factors are considered in our analysis accounting for low (CD/ISD=0.1), medium (CD/ISD=1.0) and high (CD/ISD=2.0) correlation levels. An illustration of the partially correlated patterns arising for two different ratios of CD/ISD is depicted in Figure 4.3.

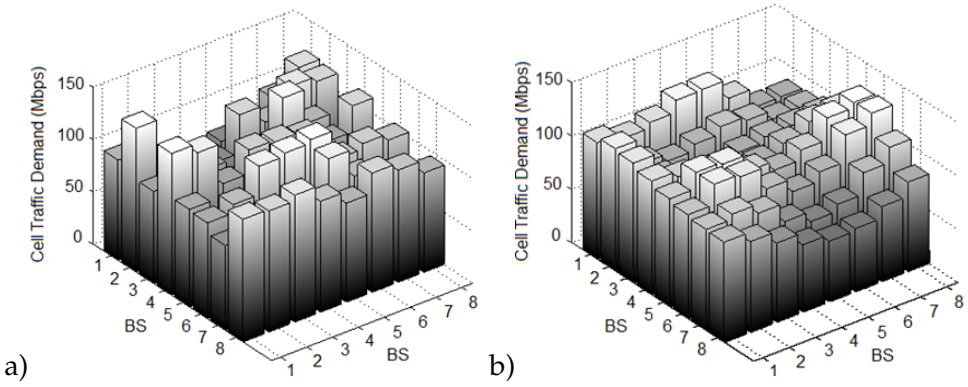

FIGURE 4.3: Traffic Demand for 64 BSs and Different Spatial Correlation Levels; (a) CD/ISD=0.1; (b) CD/ISD=1.0.

4.4 Numerical Assessments

The following assessment has been conducted to establish the underlying capacity dimensioning criteria needed for the analysis of different traffic distribution strategies. In particular, the analysis reported in this section focuses on a basic traffic distribution strategy that uses the satellite component exclusively for traffic overflow from terrestrial to satellite when the traffic demand exceeds the terrestrial satellite capacity, without discriminating between particular services. First, dimensioning results for BS and $A1$ levels in terms of over-dimensioning factors to accommodate a given percentile of the traffic demand are presented, then the obtained results in terms of capacity gains achieved when satellite traffic overflow is enabled are explained.

4.4.1 Dimensioning of the Backhaul Capacity Required at BS and A1 Levels

Table 4.1 presents the total backhaul capacity that should be available at BS level to accommodate a given percentile of the traffic demand. The capacity is given in terms of the over-dimensioning factor, which is computed as the ratio between the total backhaul capacity and the average traffic demand. Results have been obtained considering 10000 random realizations of the traffic demand and different values of σ/μ. For example, as shown in the table, the over-dimensioning factor to support the 90-th percentile of the traffic demand for $\sigma/\mu = 0.05$ is 1.31 (i.e. $1.31 \cdot \overline{T}$ = 131 Mbps). If instead the 99-th percentile of the traffic demand was to be accommodated, this would require increasing the backhaul capacity up to 1.66 (i.e. 166 Mbps), that is, 35 Mbps of additional capacity to move from the 90-th to the 99-th percentiles. This capacity increase between the 90-th and 99-th percentiles becomes considerably larger for higher values of σ/μ, e.g. 271 Mbps of additional capacity for $\sigma/\mu = 0.02$.

Similarly, as results in Table 4.1 for the dimensioning of BS links, Table 4.2 presents the over-dimensioning factor that would be required for the $A1$ links to accommodate the 90-th and 99-th percentiles of the traffic demand. Results are given for a different number of BSs that can be anchored to the same aggregation point, denoted as aggregation factor A_f, and for different σ/μ ratios. In this case, results also

TABLE 4.1: Over-Dimensioning Factor at BS Level.

σ/μ	**Percentile 75**	**Percentile 90**	**Percentile 95**	**Percentile 99**
0.02	1.06	1.12	1.16	1.23
0.05	1.14	1.31	1.42	1.66
0.1	1.21	1.61	1.87	2.64
0.2	1.25	2.06	2.85	4.77

depend on the correlation factor CD/ISD between the traffic of the BSs that is being aggregated. For the sake of easing the comparison with results in Table 4.2, the over-dimensioning factor is normalized by the aggregation factor A_f. For example, for $A_f = 8$, $\sigma/\mu = 0.05$ and low correlation settings (CD/ISD=0.1), over-dimensioning factors to satisfy the 90-th and 99-th percentiles are, respectively, 1.09 and 1.21 This means that the additional capacity required per BS to go from the 90-th to the 99-th percentile is around 12 Mbps per BS, equivalent to a total increase of 96 Mbps capacity in the $A1$ link. For higher correlation settings (CD/ISD=2.0), the additional capacity increases to 30 Mbps per BS, leading to a total of 240 Mbps capacity increase required in the $A1$ link.

TABLE 4.2: Over-Dimensioning Factor at $A1$ Level.

A_f	σ/μ**=0.02 Percentile 90/99**		σ/μ**=0.05 Percentile 90/99**		σ/μ**=0.1 Percentile 90/99**	
	CD/ISD =0.1	**CD/ISD =2.0**	**CD/ISD =0.1**	**CD/ISD =2.0**	**CD/ISD =0.1**	**CD/ISD =2.0**
2	1.08/1.16	1.10/1.21	1.21/1.43	1.29/1.65	1.43/2.02	1.59/2.45
4	1.05/1.11	1.11/1.19	1.15/1.30	1.26/1.61	1.30/1.68	1.59/2.43
8	1.04/1.08	1.08/1.16	1.09/1.21	1.28/1.58	1.20/1.46	1.54/2.35
16	1.02/1.05	1.05/1.13	1.07/1.15	1.26/1.54	1.13/1.31	1.53/2.24
32	1.01/1.04	1.04/1.09	1.04/1.11	1.21/1.44	1.09/1.24	1.42/1.99

4.4.2 Capacity Gain at BS Level

This section assesses the trunking gain that can be achieved when satellite traffic overflow is enabled at BS level. The idea is that the overall backhaul capacity at BS level (terrestrial and satellite) shall be able to accommodate the 99-th percentile of the total traffic demand of the BS. To that end, the terrestrial capacity is assumed to be dimensioned to cope with the 90-th of the demand while the satellite capacity shall cope with the excess of the demand between the 90-th and 99-th percentiles. Under

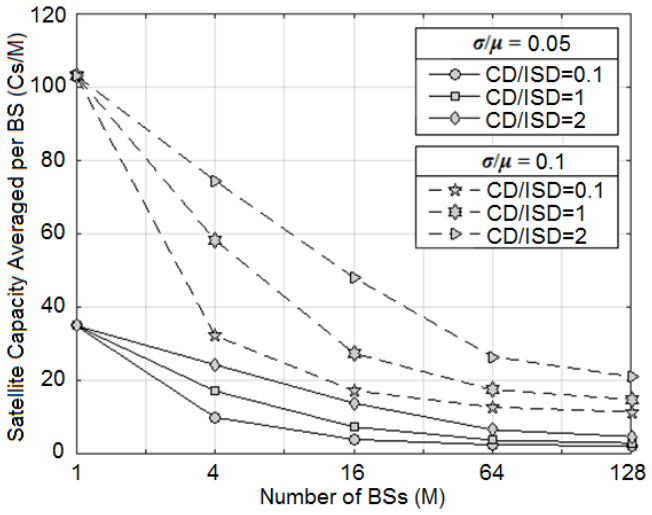

FIGURE 4.4: Satellite Capacity Averaged per BS (C_S/M) that is Required for Satellite Traffic Overflow. It is Calculated for $CD/ISD = [0.1, 1, 2]$, $\sigma/\mu = [0.05, 0.1]$ and a Terrestrial Link Capacity Dimensioned for the 90-th Percentile.

these assumptions, Figure 4.4 shows the amount of total satellite capacity averaged per BS (C_S/M) that would be required for different values of M, correlation factors CD/ISD and σ/μ ratios. As shown in the figure, if satellite capacity was allocated per BS in a dedicated manner $(M = 1)$, a total of 35 Mbps would be required per BS for $\sigma/\mu = 0.05$.

On the other hand, if the same satellite capacity is shared among $M = 128$ BSs, only 3.1 Mbps shall be allocated on average per BS for medium correlation levels $(CD/ISD = 1.0)$ and $\sigma/\mu = 0.05$, yielding a capacity reduction by a factor of 11. If instead $\sigma/\mu = 0.1$ is considered, a total of 103 Mbps is required per BS for $M = 1$, dropping to 15 Mbps per BS for $M = 128$, which is equivalent to a capacity reduction by a factor of 6.8. As expected, this capacity reduction factor varies with the degree of spatial correlation. As shown in Figure 4.4, the higher the spatial correlation, the greater the amount of capacity that is required to cope with the traffic overflow, because the traffic peaks tend to appear more simultaneously. However, even under high correlation levels $(CD/ISD = 2.0)$, capacity reduction factors between 7.2 and 4.8 are achieved for $M = 128$ and $\sigma/\mu = 0.05$ and $\sigma/\mu = 0.1$, respectively. On the other hand, grouping BSs with low correlation levels leads to capacity reduction factors between 15 and 8.9.

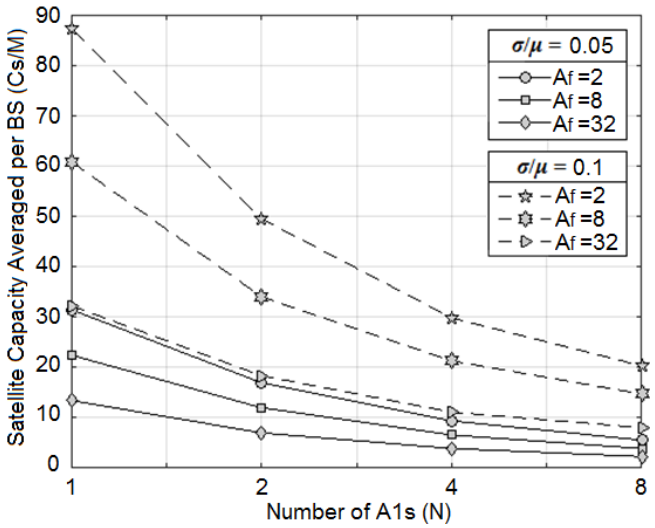

FIGURE 4.5: Satellite Average Capacity per BS (C_S/M) that is Required for the 90-th Percentile of the Satellite Traffic Overflow at $A1$ Level with Different Aggregation Factors (A_f). It is Calculated for $CD/ISD = 1$, $\sigma/\mu = [0.05, 0.1]$ and a Terrestrial Link Capacity Dimensioned for the 90-th Percentile.

4.4.3 Capacity Gain at A1 Level

This section assesses the achievable trunking capacity gain if the satellite capacity is deployed at $A1$ level, in contrast to the previous case where satellite traffic overflow was assumed to be possible at each individual BS site. The results in Figure 4.5 allow us to assess the trunking gain that can be achieved by sharing some amount of satellite capacity for traffic overflow among a number of $A1$ aggregation points.

Similarly to the previous analysis for BS links, the terrestrial backhaul capacity of the $A1$ links is dimensioned to absorb the 90-th percentile and the shared satellite capacity is used to achieve the 99-th percentile. From the provided results, focusing on the example given for $A_f = 8$ and $\sigma/\mu = 0.05$, if the satellite capacity is shared among 4 $A1$ aggregation points, the total amount of required satellite capacity is $7 \cdot 4 \cdot 8$ (i.e. 224 Mbps), equivalent to 56 Mbps of additional capacity per aggregation point. This results in over a 3-fold reduction in comparison with the 192 Mbps of additional capacity required if the satellite capacity was just dedicated to a single $A1$. For the other cases shown in Figure 4.5, the capacity reduction factor varies between 4 and 6. As expected, these capacity reduction factors at $A1$ level are lower than those achieved when satellite traffic overflow is supported at BS level, but still meaningfully in the case of $A1$ points with low aggregation factors.

4.5 Summary

This chapter has developed a general model and problem formulation for the analysis of capacity and traffic management strategies in hybrid satellite-terrestrial mobile backhauling networks that rely on SDN for traffic steering at flow level.

An initial assessment has been reported to estimate the capacity gains that the satellite component could bring when used to support a basic traffic overflow strategy. The results obtained have revealed that, besides further motivations supporting the interest of the satellite backhaul (e.g. service extension in hard to reach areas, increased resilience, fast and temporary deployments), relevant capacity gains can be achieved thanks to the multiplexing gain of sharing some amount of satellite capacity across a number of BS sites and/or $A1$ traffic aggregation points. In particular, it has been shown under a realistic traffic model that if the shared satellite capacity is deployed to accommodate the traffic demand between the 90-th and 99-th percentiles, capacity reduction factors within a range of 5-15 and 4-6 can be achieved with respect to the case that such capacity is provisioned in a dedicated manner at BS or $A1$ levels, respectively.

In the subsequent chapters are considered the formulation and analysis of more sophisticated traffic steering strategies with QoS differentiation and the consideration of multicast traffic.

Chapter 5

Network Utility Maximization Framework and Performance Evaluation of TE Strategies

5.1 Introduction

Taking into account the previous hybrid satellite-terrestrial backhaul network capacity dimensioning results, this chapter focuses on the formulation of an SDN-based TE application that exploits the allocation criteria capacity depending on the traffic nature; admission control and rate control features. To that end, it is built a utility framework model for the analysis of traffic distribution strategies, seeking the network utility maximization criteria.

The new SDN-based TE application considers a hybrid satellite-terrestrial mobile backhaul network scenario where a satellite component is used to complement the terrestrial infrastructure in a way that end-to-end paths across both satellite and terrestrial links can be centrally computed and re-arranged dynamically at flow-level granularity in front of link congestion and failure events, and is designed to cope with a satellite capacity provisioned to improve the resilience of a hybrid satellite-terrestrial mobile backhaul network. Unlike more basic strategies that might be devised for simply replacing a failed terrestrial link with satellite capacity or just activating traffic overflowing through satellite in high demanding peak-times, the proposed scheme pursues an optimal allocation of the available satellite and terrestrial capacity so that a network utility is maximized under both failure and non-failure

terrestrial links conditions.

The rest of the chapter is organized as follows; first, the previous works related to the subject are briefly presented and explained; subsequently, it is provided and descripted the system model and the formulation of the problem for optimal traffic distribution; Given the NP-hard nature of the resulting optimization problem, it is descripted the heuristic we have relied on for the assessment; Finally, a performance analysis is presented to assess the behavior of the proposed SDN-based TE application under diverse scenarios, including terrestrial link failures in some BSs and a number of transportable BSs that exclusively rely on the satellite capacity for backhauling.

5.2 Related Work

Remarkably, the roles and benefits of satellite networks in 5G have been introduced and discussed in 3GPP, with several use cases identified in many study items (e.g. support of, "5G connectivity via satellite" within 3GPP TR23.799, the "higher availability" requirement within 3GPP TR22.862, etc.). As a result, a requirement for 3GPP systems to be able to provide services using satellite access has been included within the normative Stage 1 requirements for next generation mobile telecommunications being elaborated by 3GPP [9]. This is especially of interest in public protection and disaster relief (PPDR) efforts given the high dependability on communication systems for effective disaster mitigation [11], especially communications in remote/rural areas that might require the fast deployment of network capacity as well as in distressed areas where the terrestrial backhaul infrastructure might have suffered damages. One of the ways to minimize the network vulnerabilities in case of a disaster is introducing an appropriate redundancy within the network [11]. In this case, satellite capacity can be deployed as a redundant backhaul capacity to any network node and also may be available to operate in challenging post disaster scenarios, allowing rapid emergency communication network deployments as those based on vehicular or transportable network nodes (referred to as mobile cells) described in [11], e.g., base stations (BS). Satellite links could provide additional bandwidth to backup connectivity to critical cell sites as well as to divert traffic from

congested areas so that the capacity in the terrestrial links could be supplemented during peak-times or even replaced in case of total/partial failures as well as for emergency mobile cell deployments.

However, despite the potential value of using a satellite component for resilience purposes in mobile networks is generally recognized [2][12], to the best of authors' knowledge, there is a lack of published studies focused on the development of traffic distribution strategies aimed to improve the hybrid satellite-terrestrial backhaul network performance or just assessing the benefits in terms of that a satellite capacity deployed for resilience purposes can bring into a hybrid terrestrial-satellite backhauling network scenario. Only a few related works coping with network design and traffic engineering in hybrid satellite-terrestrial backhaul networks are available in the open scientific literature, though the applicability area is not specifically that of resilience. In particular, the works in [81]-[83] are aimed at exploiting the wide coverage capability of the satellite component for broadcast/multicast applications, showing that satellite links can relieve a Long Term Evolution (LTE) mobile network of a significant part of the multimedia broadcast multicast services traffic. Under this application, traffic distribution schemes are mainly intended to establish the best routes for multicast connections (i.e. multicast routing problem) so that the multicast performance is enhanced by seeking e.g. the minimal routing end-to-end tree delay, the minimum tree cost or other minimum failure ratio [82][84] of multicasting connections, or the combination of them [85]. Some examples of heuristic algorithms proposed in this context are given in [82][86][87]. Another addressed application area is that of emergency communications. In this context, a limited satellite capacity is typically deployed to replace terrestrial backhaul links where these are not available (e.g. terrestrial links disrupted by natural disasters) and prioritization mechanisms are applied to manage the satellite connectivity. For example, authors in [88] propose a scheme for service prioritization under traffic congestion based on the communication needs of first responders under different emergency scenarios. In a related work, [89] studies a service classification and management scheme with two traffic classes: streaming and background. The streaming traffic class gets higher priority and ensures that the constant data rate is available to the user and the rest of the satellite channel capacity is assigned to the background

class where the available bit rate may vary. Furthermore, the issue of resilience and congestion in hybrid satellite-terrestrial wireless backhaul networks is also being researched in [90] through the implementation of smart antennas for dynamic network topology reconfiguration according to traffic demands.

5.3 Network Model

Let's consider a cellular network with M BSs deployed across a large geographical area at fixed locations referred to as Fixed Base Stations (FBS) where the transport connectivity between the M FBSs and the mobile core network is assumed to be delivered through a hybrid satellite-terrestrial backhaul network. In particular, we assume that each FBS is connected to the transport infrastructure through a terrestrial link (e.g., microwave or wired link) and that there is also a ST co-located at each site to provide satellite connectivity.

The maximum terrestrial and satellite link capacity at the m-th fixed base station (FBS_m) is denoted, respectively, as $C_{FBS-T,m}$ and $C_{FBS-S,m}$. Accordingly, the maximum terrestrial and satellite capacity available at the M FBS are represented, respectively, by the capacity vectors $\mathbf{c}_{FBS-T} = \{C_{FBS-T,1}, ..., C_{FBS-T,M}\}$ and $\mathbf{c}_{FBS-S} = \{C_{FBS-S,1}, ..., C_{FBS-S,M}\}$. In addition, it is assumed that maximum aggregate satellite capacity in use at a given time across all FBSs cannot exceed C_S. Moreover, the terrestrial link availability is captured through a binary vector $\mathbf{a} = \{a_1, ..., a_M\}$ where $a_m = 1$ stands for the terrestrial link being operational and $a_m = 0$ represents a Link Down (LD) situation. An illustration of the network model is depicted in Figure 5.1.

5.4 Traffic Model

Our analysis considers mixes of unicast elastic and unicast inelastic traffic. Inelastic/stream traffic is generated by time-sensitive applications, like e.g. Voice over IP (VoIP) and Video Streaming on Demand (VSOD) that typically has strict bandwidth and/or delay requirements. Elastic traffic on the other hand is generated by applications such as web browsing and file-transfers where the delivered bit rate and/or the download time are more important than inter-packet or end-to-end delays. Indeed,

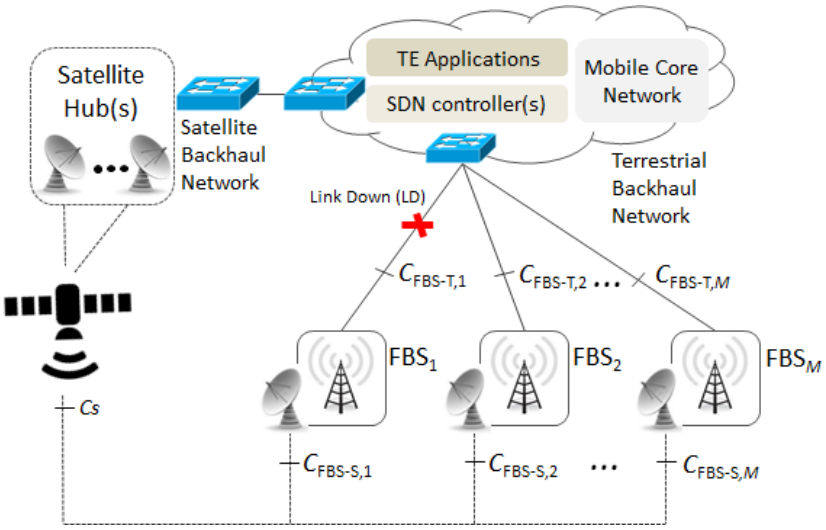

FIGURE 5.1: Hybrid Satellite-Terrestrial Backhaul Network Model.

this traffic classification is captured in the QoS model established for LTE systems by considering two types of bearer services that can be enforced in the network: Guaranteed Bit Rate bearers (GBR) and Non-GBR bearers. Thus, the unicast traffic flows served through GBR bearers (called UG flows in the following) are given a minimum guaranteed bit rate to operate satisfactorily; otherwise, the quality might be severely affected. On the other hand, the unicast traffic flows served through Non-GBR bearers (called UN flows in the following) do not get such a minimum bit rate reservation but can see a wide variability of the achieved bit rate, being more exposed to congestion related packet losses and/or delay variability (without necessarily a noticeable impact on QoS).

Let's define $I_S(m)$ as the set of service flows types $s \in \{UG, UN\}$ served through the FBS_m. Then, let $x_{s,m,i}$ and $r_{s,m,i}$ (denoted as matrices X and R respectively) denote respectively, the link selection (if the flow is served through the satellite backhaul link $[x = 0]$ or through the terrestrial backhaul link $[x = 1]$) and bit rate allocation variables associated with the i-th flow of s service type at the FBS_m.

5.5 Utility Framework

The conception of the TE logic requires that a specific traffic and link characterization is first established. This is necessary to determine, if applicable, the QoS requirements, such as maximum tolerable latency and jitter, minimum required bandwidth, etc. per type of service/user that need to be fulfilled in order to achieve a given QoE/satisfaction level. To that end, we resort to the use of utility functions to describe the QoE/satisfaction level that is achieved when a particular flow is served across the hybrid satellite-terrestrial backhaul. In our case, utility functions are formulated to account for two main aspects; (1) the bit rate of the flow that can be allocated across the E2E path and; (2) whether the E2E path traverses a satellite link or not (i.e. the higher delay incurred when using a satellite link can result in some level of service degradation that is reflected with a lower utility).

The formulation of the utility functions is also dependent on the nature of the services. Based on the above considerations, the utility functions considered in our analysis for the characterization of UG and UN services are provided in Figure 5.2. The utility functions account for the delivered bit rate (r) and consider whether the flow is served through the satellite or terrestrial backhaul link (x).

Commonly, the step function is used for UG flows [91]. In particular, in our analysis a two-level step function is used for UG flows, reflecting two possible bit rates/quality levels that could be on offer (e.g. standard and high definition VSODs). This UG utility function, defined by Eq. 5.1 - 5.3, is parameterized by the bit rates R_1^{UG} and R_2^{UG} to be delivered for the standard/high quality offerings, respectively, a utility reduction factor p^{UG} to account for the potential quality/satisfaction degradation when using satellite links instead of terrestrial one and a utility reduction factor α^{UG} to account for the impact of rate selection between R_1^{UG} and R_2^{UG} as shown in Figure 5.2 (a).

The UG utility function is:

$$U^{UG}(r,x) = U_o^{UG}(x) \cdot U_r^{UG}(r) \tag{5.1}$$

where

$$U_o^{UG}(x) = p^{UG} + x\left(1 - p^{UG}\right) \tag{5.2}$$

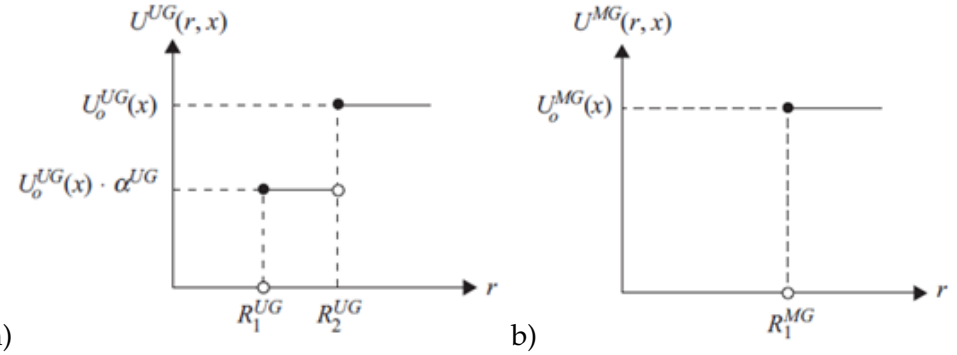

FIGURE 5.2: a) UG Utility Function, and b) MG Utility Function.

and

$$U_r^{UG}(r) = \begin{cases} 0 & 0 < r < R_1^{UG} \\ \alpha^{UG} & R_1^{UG} < r < R_2^{UG} \\ 1 & r \geq R_2^{UG} \end{cases} \tag{5.3}$$

Although in this chapter we will not consider multicast GBR services (called MG services in the following), in the later chapter we will consider it in our analysis, so we advance here the utility function utilized for this kind of traffic. For the characterization of MG service flows, a one-level step utility function is used, as defined by Eq. 5.4 - 5.6. In this case, R_1^{MG} is the minimum bit rate to be delivered for the high quality and the parameter p^{MG} is a utility reduction factor to account for the potential quality/satisfaction degradation when using satellite links instead of terrestrial as shown in Figure 5.2 (b).

The MG utility function is:

$$U^{MG}(r,x) = U_o^{MG}(x) \cdot U_r^{MG}(r) \tag{5.4}$$

where

$$U_o^{MG}(x) = p^{MG} + x\left(1 - p^{MG}\right) \tag{5.5}$$

and

$$U_r^{MG}(r) = \begin{cases} 0 & 0 < r < R_1^{MG} \\ 1 & r \geq R_1^{MG} \end{cases} \tag{5.6}$$

With regard to UN service flows, the utility functions can be more diverse [92], depending on which specific aspects/service characteristics that is wanted to stress.

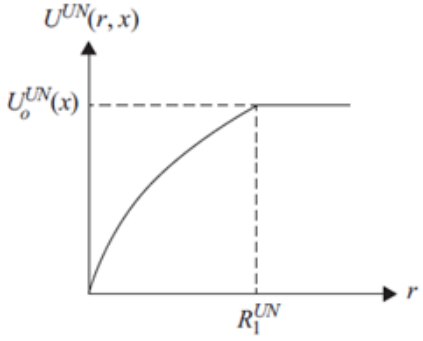

FIGURE 5.3: UN Utility Function.

In our case we have adopted a logarithmic utility function [93], which is one of the most commonly used and already serves our needs. The normalized utility function of UN service flows is defined by Eq. 5.7 - 5.9, where R_1^{UN} is used to establish the bit rate for which is considered that the service is already provided with a good quality (so no utility gain is envisioned by serving UN service flows with higher bit rates) and the parameter p^{UN} is the utility reduction factor for UN service flows. For UN service, the utility functions are represented by Figure 5.3.

The UN utility function is

$$U^{UN}(r,x) = U_o^{UN}(x) \cdot U_r^{UN}(r) \tag{5.7}$$

where

$$U_o^{UN}(x) = p^{UN} + x\left(1 - p^{UN}\right) \tag{5.8}$$

and

$$U_r^{UN}(r) = \begin{cases} \frac{\log(r+1)}{\log\left(R_1^{UN}+1\right)} & if \quad 0 < r < R_1^{UN} \\ 1 & if \quad r > R_1^{UN} \end{cases} \tag{5.9}$$

5.6 TE for Optimal Traffic Distribution

Different resilience schemes can be defined based on a satellite capacity allocated to cope with the failure of terrestrial links and how this capacity is intended to be used. In particular, in our case a resilience scheme is characterized by a ratio $N : M$ that

indicates that the satellite capacity has been dimensioned to cope with the failure of N terrestrial links in a group of M BSs. On this basis, given the maximum transport capacity deployed at BSs sites ($\mathbf{c}_{FBS-T}$ and $\mathbf{c}_{FBS-S}$), the adopted resilience scheme $N : M$ (from which the maximum aggregate satellite capacity C_S is derived), the terrestrial link availability ($\mathbf{a}$), and the traffic demand t^{UG} and t^{UN}, the following optimization problem can be formulated for finding the best traffic distribution in terms of achievable bit rates and use of the terrestrial or satellite capacity per connection, represented respectively by matrices R and X ($r_{s,m,i}$ and $x_{s,m,i}$):

Given: F_{BS}(Set of Fixed BSs)

$\mathbf{c}_{FBS-T}$, $\mathbf{c}_{FBS-S}$, C_S, $\mathbf{a}$ (Link capacity vectors, satellite system capacity and availability vector, respectively)

$I_s(m)$, (Set of service flows types)

Find: $\{x_{s,m,i}\}$,$\{r_{s,m,i}\}$ (Link Selection and Bit Rate)

Max:

$$\sum_{s\in\{UG,UN\}}\sum_{m=1}^{M}\sum_{i\in I_s(m)} U^S(r_{s,m,i}, x_{s,m,i}) \tag{5.10}$$

s.t.

$$\sum_{s\in\{UG,UN\}}\sum_{i\in I_s(m)} (r_{s,m,i}\cdot x_{s,m,i}) \leq a_m \cdot C_{FBS-T,m} \qquad \forall m \tag{5.11}$$

$$\sum_{s\in\{UG,UN\}}\sum_{i\in I_s(m)} (r_{s,m,i}\cdot (1-x_{s,m,i})) \leq C_{FBS-S,m} \qquad \forall m \tag{5.12}$$

$$\sum_{s\in\{UG,UN\}}\sum_{m=1}^{N+M}\sum_{i\in I_s(m)} (r_{s,m,i}\cdot (1-x_{s,m,i})) \leq C_S \tag{5.13}$$

$$x_{s,m,i} \in \{0,1\}, r_{s,m,i} \in \{0, \Re^+\} \qquad \forall s,m,i \tag{5.14}$$

where Eq. 5.10 is the objective function defined as the sum of service flows utilities, Eq. 5.11 - 5.13 set out, respectively, the capacity constraints of terrestrial links per BS, satellite links per BS and maximum aggregated satellite capacity and Eq. 5.14 restricts the the possible values of the decision variables. The resulting problem is a non-linear optimization problem that falls within the category of non-convex mixed-integer nonlinear programming, which is known to be NP-hard [94].

5.6.1 Traffic Distribution Strategy Algorithm

Given the NP-hard nature of the optimization problem formulated in the previous section, we have developed an heuristic algorithm that can handle, with low complexity, large scenarios in terms of number of BSs and traffic flows. The proposed algorithm seeks to find a feasible solution to the optimization problem by splitting it into two separate sub-problems:

- First, the bit rate allocation and path selection through either terrestrial or satellite backhaul is computed for the UG flows/connections. This is done on a connection–basis trying to allocate, per BS and sequentially, the link and bit rate that provides the highest utility per connection. When capacity contraints are reached, the algorithm attempts to allocate pending connections by gradually reducing the utility of the new as well as of the pre-established connections as long as the global utility (i.e. term $U^{UG}(R;X)$ in Eq. 5.10) can be still increased. The procedure is detailed in Algorithm 1.

- Second, the remaining satellite capacity is distributed seeking a max-min rate allocation for UN connections irrespective of using satellite or terrestrial capacity. It's worth noting that a max-min rate allocation is equivalent to the maximization of the global utility of UN traffic (i.e. term $U^{UN}(R;X)$ in Eq. 5.10) when there is no distinction between using satellite or terrestrial capacity [95], which in our case is strictly valid for factor $p^{UN} = 1$. This is carried out by first distributing the satellite capacity among the BSs in proportion to the number of UG connections handled per BS so that overall capacity left for UG services is balanced across BSs. Then, in each BS, the satellite capacity is equally shared among the served connections. The detailed procedure is given in Algorithm 2.

Based on above considerations, we have solved the network utility maximization problem by splitting it in two separate sub-problems. In particular, the problem is solved sequentially first considering only UG services (sub-problem #1), then adding the UN services (sub-problem #2). Some details on the algorithms used to solve each of the sub-problems are provided in the following.

Sub-problem #1 can be indeed formulated as the general problem in Eq. 5.10 - 5.14 with $s \in \{UG\}$. This sub-problem is a discrete optimization problem (note that utility function for UG traffic is a step function) that is solved by means of a heuristic. The proposed heuristic is based on a greedy algorithm, which, flow by flow, decides on the link and bit rate that yields the highest utility. In this way, as long as capacity constraints are met, the algorithm always moves towards getting the maximum utility. When capacity constraints are found and a flow cannot be accepted with its maximum utility, the algorithm revisits some of previous decisions in order to see if some modifications are possible (e.g. rate reduction or link change) as long as these modifications result into higher network utility. The algorithm is detailed in Algorithm 1. For the allocation and re-allocation decisions, the algorithm uses a set of precomputed utility increments ΔU_n that account for the potential combinations of link and rate selection of the current and up to q previous allocated flows. In the case of sub-problem #1 for UG services, these increments are computed as follows: $\Delta U_1 \leftarrow U^{UG}(R_2^{UG},1)$, $\Delta U_2 \leftarrow U^{UG}(R_2^{UG},0)$, $\Delta U_3 \leftarrow U^{UG}(R_1^{UG},1)$ and $\Delta U_4 \leftarrow U^{UG}(R_1^{UG},0)$, for allocations that do not impact on previous flows and $\Delta U_5 \leftarrow 2U^{UG}(R_1^{UG},1)$- $U^{UG}(R_2^{UG},1)$, $\Delta U_6 \leftarrow 2U^{UG}(R_1^{UG},0)$- $U^{UG}(R_2,0)$,..., $\Delta U_{n-1} \leftarrow (q+1)U^{UG}(R_1^{UG},1)$- $qU^{UG}(R_2^{UG},1)$, $\Delta U_n \leftarrow (q+1)U^{UG}(R_1^{UG},0)$-$qU^{UG}(R_2^{UG},0)$, for allocations that force a rate reduction in previously allocated flows.

Procedure: Resource Allocation (service type (s), S_{BS}**,** $\in \{FBS\}$***)***

$\Delta U = \{\Delta U_1, \Delta U_2, ..., \Delta U_n\}$, Set of utility increments different potential allocations

$r_{s,m,i} \leftarrow 0$ for all *m,i*

for $FBS_m \in S_{BS}$ **do**

  $AvailableCapacity = TRUE$

  for $i \in I_s(m)$ *and while Available Capacity=TRUE* **do**

    Select $(r_{s,m,i}, x_{s,m,i})$ and the applicable rate deruction of previous allocated flows yielding the highes $\Delta U_n \in \Delta U$ that satisfies Eq.5.10-5.14, if $(r_{s,m,i} = 0)$, $AvailableCapacity = FALSE$

  end

end

Algorithm 1: Resource Allocation Algorithm for UG service flows.

Sub-problem #2 for utility maximization of UN services can also formulated from Eq. 5.10-5.14 with $s \in \{UN\}$ and the following new link capacities.

$$C'_{FBS-T,m} = C_{FBS-T,m} - \sum_{s\in\{UG\}} \sum_{i\in I_s(m)} r_{s,m,i} \qquad \forall m \tag{5.15}$$

$$C'_{FBS-S,m} = C_{FBS-S,m} - \sum_{s\in\{UG\}} \sum_{i\in I_s(m)} r_{s,m,i} \qquad \forall m \tag{5.16}$$

$$C'_S = C_S - \sum_{m=1}^{M} \sum_{s\in\{UG\}} \sum_{i\in I_s(m)} r_{s,m,i} \tag{5.17}$$

In this case, the sub-problem turns to be an optimization problem which combines a concave utility function with a discrete link selection variable, with constraint functions that contain products of both concave and discrete optimization variables. Despite the problem shows complexity regarding the application of a direct method for its solution, the consideration of a single type of traffic $s \in \{UN\}$ facilitates its analysis and the introduction of a heuristic. Indeed, in [95] it is demonstrated that utility proportional fair resource allocation ensures utility max-min fairness for all users sharing a single path in the network, i.e. the utility that should be shared fairly among users. Then, a max-min rate allocation is equivalent to the maximization of the utility at each link for UN traffic. This same concept can be extended to the global utility of the network when there is no distinction between using satellite or terrestrial capacity, which in our case is strictly valid for factor $p^{UN} = 1$. Since in the network model a service flow can be handled only by a single link, and in turn each BS has at most two possible links, for $p^{UN} = 1$, a heuristic can find the utility maximization applying the above mentioned rule, selecting the link to each UN flow in such a way that all service flows have as much as possible the same bit rate utilizing the full links capacities. To do this, the algorithm solve the number of UN flows handled by terrestrial or satellite links at each FBS_m(K_m^T and K_m^S respectively) that theoretically would give the bit rate as similar as possible to all UN connections by solving the Eq. 5.15 and 5.16 subject to Eq. 5.11-5.14:

$$r_{UN,m,i} = \frac{C'_{FBS-T,m}}{K_m^T} = \frac{C'_{FBS-S,m}}{K_m^S}, \quad \forall i : i \in I_{UN}(m), \quad m = 1, ..., M \tag{5.18}$$

where

$$r_{UN,1,i} = r_{UN,2,i} = \cdots = r_{UN,M,i}, \qquad \forall i \tag{5.19}$$

s.t. Eq. 5.11 and 5.14.

where K_m is defined as the total of UN service flows at the FBS_m. For $p^{UN} < 1$, besides solving Eq. 5.18 and 5.19 and fixing the satellite links utilized satellite capacities $C'_{FBS-S,m}$, the algorithm reselects the right links $x_{UN,m,i} \forall i \in \{I_{UN}(m)\}$, in order to get K_m^T and K_m^S that maximize the sum of UN service flows utilities at each BS_m. The procedure is detailed in Algorithm 2.

Procedure: Resource Allocation (UN flows, $S_{BS} \in \{FBS\}$)

$r_{UN,m,i} \leftarrow 0$ for all m, i

for $FBS_m \in S_{BS}$ **do**

  K_m^T,K_m^S,$C'_{FBS-S,m} \leftarrow$ Solve Eq. 5.18 - 5.19;

end

for $FBS_m \in S_{BS}$ **do**

  for $i \in \{I_{UN}(m)\}$ *and while Available Capacity=TRUE* **do**

    Given $C'_{FBS-S,m}$, find (K_m^T, K_m^S) that yields the highest $\Delta U_n \in \Delta U$ that satisfies Eq.5.11 - 5.14, this is,

    $\{x_{UN,m,i}\}$,$\{r_{UN,m,i}\}$,K_m^T,$K_m^S \leftarrow argmax U(x_{UN,m,i}, r_{UN,m,i})$;

  end

end

Algorithm 2: Resource Allocation Algorithm for UN Service Flows.

5.7 Numerical Assessment

5.7.1 Simulation Method and Scenario Settings

The numerical assessment is based on a Monte Carlo simulation method that solves the traffic distribution problem in a set of 1000 realizations. In each realization, a

TABLE 5.1: Scenario Settings for the Numerical Assessment.

Parameter	Values
Number of FBSs (M)	16
Terrestrial link capacity $(C_{FBS-T,m})$	131 Mbps
Terrestrial link availability $(\mathbf{a})$	0-4 Link Down
Maximum satellite link capacity per $FBS_m(C_{FBS-S,m})$	200 Mbps
Resilience scheme	RS, 1:16, 1:18, 1:4
Maximum aggregate satellite capacity (C_S)	0, 131, 262, 524 Mbps
Average number of UG connection per FBS_m	0-50
Average number of UN connection per FBS_m	0-50
Standard quality rate for UG connections (R_1^{UG})	3 Mbps[a]
High quality rate for UG connections (R_2^{UG})	6 Mbps[a]*a*
Utility reduction factor due to UG link selection (α^{UG})	0.6
Maximum utility rate UN (R_1^{UN})	13 Mbps[b]
Utility reduction factor over satellite (p^{UG} and p^{UN})	0.6, 0.8, 1

[a]Typical mobile Video Resolution and Bitrates [97].
[b]The global average for LTE download speeds (source: "The State of LTE", OpenSignal, February 2016) [98].

random number of UG and UN connections is generated per BS according to a uniform distribution. Table 5.1 provides the range of values considered for the different model parameters in the numerical assessment. Without any loss of generality and for the sake of consistency, the values considered for service characterization as well as backhaul capacities are inspired in current state-of-the-art 4G and satellite broadband technologies.

With regard to the capacity of the terrestrial links, the considered setting (131 Mbps) is based on the dimensioning analysis presented in previos section, to cope with the 90-th percentile of the traffic demand when considering a realistic traffic model that exhibits a log-normal distribution with an average load of 100 Mbps per BS. This value is then considered to establish the range of values for the maximum aggregate satellite capacity (C_S) according to the adopted resilience scheme. On the other hand, the maximum satellite link capacity per $FBS_m(C_{FBS-S,m})$ is set to 210 Mbps, in line with today's top-of-the-line satellite modems based on DVB-S2X [96].

5.7.2 Numerical Assessment Under no Link Failure Conditions

This first assessment is intended to show the performance of the traffic distribution strategy when all terrestrial links are operational (i.e. LD=0). Unlike a more classical

approach where the satellite capacity is only used as failover/back-up, the idea here is to see how the proposed strategy can get the most utility out of all the installed capacity.

Focusing first on the UG traffic performance, the achieved average utility per UG connection when no satellite capacity for resilience is deployed (NR case) is given in Figure 5.4, along with the utility gain achieved when considering the different resilience schemes (RS1:16, RS1:8 and RS1:4) under different UG traffic loads. It is observed that, under NR, up to around 15 UG connections can be served on average per BS while achieving the maximum utility. From that point, utility starts to decrease and it's when the satellite capacity gets into play to deliver the amount of utility gain depicted on the figures on the right-hand axis. In particular, achieved utility gains for RS=1:4 sit between 8% (considering a factor $p^{UG} = 0.6$) and 18% (considering a factor $p^{UG} = 1$) when the number of connections goes between 25-30, therefore raising the utility achieved from 80% under NR to levels between 88% and 98% under RS=1:4. Similar trends are obtained for UN traffic, as illustrated in Figure 5.5, noting however that in this case the utility increase can even reach higher values.

The impact on the allocated bit rates is depicted in Figure 5.6. For UG connections, the availability of satellite capacity makes that average number of UG connections per BS that can be served at full rate increases from 15 to 26 connections (70% increase) for RS1:4. For UN connections, Figure 5.6 (right) depicts the CDF of the bit rates allocated to UN traffic under different UN loads (5, 15 and 25) and considering 5 UG connections per BS and $p^{UN} = 0.8$. It can be seen as the use of a resilience scheme with higher satellite capacity does not only achieves the highest rates but also allows for a more equitable bit rate distribution among the network connections, reducing the gap of rates assigned among the UN connections. For example, if we get the UN rates within the percentiles 10% - 90%, for 15 UN connections, we get a gap of only 2 Mbps for RS1:4, while we get the gap of 3.7 Mbps and 4.5 Mbps for RS1:16 and NR respectively, this is, a reduction of almost 50% of the gap of rates when is used the resilience scheme with greatest capacity. This is due to the satellite capacity is distributed dynamically in proportion to the load on each BS, thus maintaining equitably the data rates allocated among all connections.

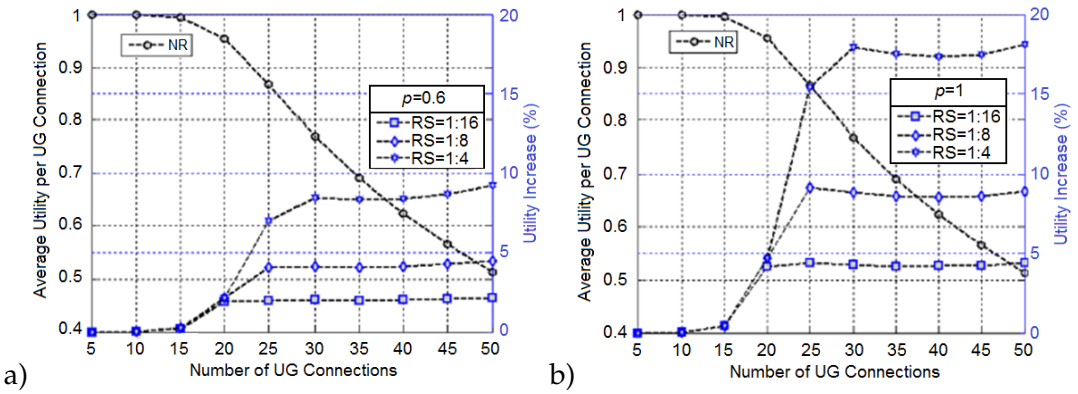

FIGURE 5.4: (a) Average Utility per UG Connection and (b) Utility Increase

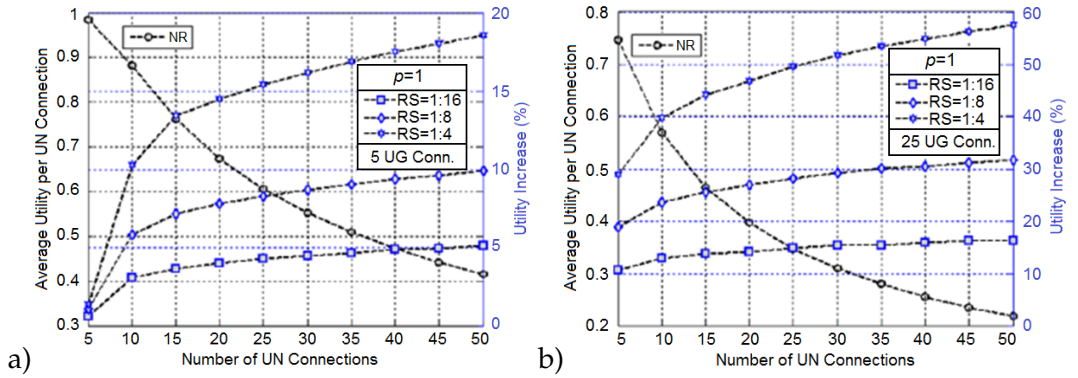

FIGURE 5.5: Average Utility per UN Connection and Utility Increase for Two UG Loads; (a) 5 and ; (b) 25 UG Connections per BS.

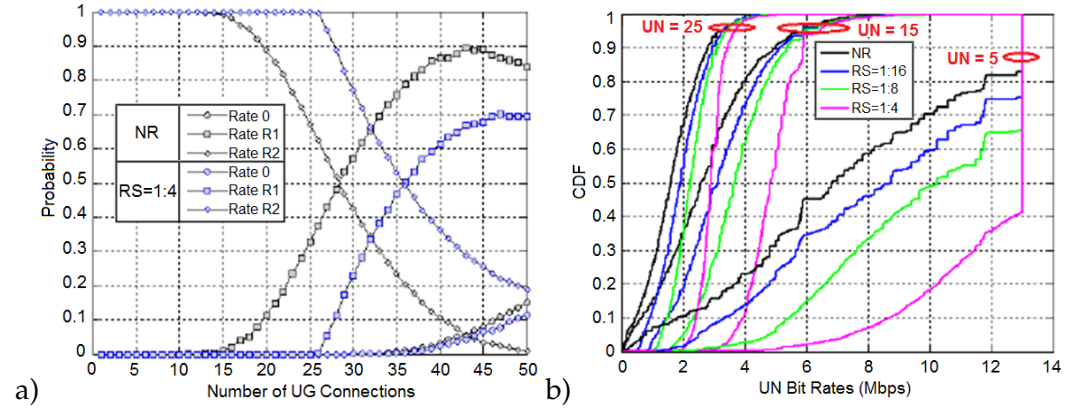

FIGURE 5.6: (a) PDFs of UG Bit Rates for NR and RS=1:4. (b) CDF of Bit Rates Assigned to UN Connections.

5.7.3 Numerical Assessment under Link Failure Conditions

Figure 5.7 (a) provides the average utility decrease per UG connection achieved in the BSs that are affected by terrestrial link failures (LD=2) in comparison with the case without link failures. Under low traffic load conditions, it can be observed that the minimum utility decrease is lower bounded by the value of p^{UG}, meaning that all the traffic served by the impaired BSs can be successfuly diverted through the satellite capacity. This utility decrease can be kept at this minimum value up to an average of 10 connections per BS when considering the resilience scheme with smaller satellite capacity (RS=1:16) and up to 24 connections for the resilience scheme with higher capacity (RS=1:4). Therefore, while the utility decrease would be of 100% for the NR case (i.e. no traffic can be served through the affected BS), the use of the satellite capacity can keep the utility reduction to values below 30% (i.e. overall utility for UG traffic above 70%) for a traffic load ranging from 20 to 30 UG connections per BS. For UN traffic, Figure 5.1 (b) shows that the impact in utility reduction is higher than that experienced by UG connections due to preferential treatment given by the traffic distribution algorithm. Moreover, additional results not included in the presented figures, show that there is also some utility reduction in the BS not directly affected by the terrestrial link failures, even though this reduction is below 4% for UG and UN connections compared to the case that all links were up.

Finally, Figure 5.8 presents the impact of the link failures in terms of the allocated bit rate distribution. For UG traffic, the average number of UG connections that can be still served at full rate increases from 7 to 23 (200% increase) when comparing RS1:4 and RS1:16, as depicted in Figure 5.8 (a). Likewise, the average number of UG connections per BS, suported with at least the minimum rate R_1^{UG}, is above 100% higher for RS1:4 than for RS1:16 (from 10 to 22 connections on average per BS). With regard to UN bit rates, Figure 5.8 (b) provides the CDF of the allocated bit rates for LD=2 and RS1:16, 1:8 and 1:4, showing the relative performance of these schemes against the a worst performing case (LD=2 and NR) and a best performing case (LD=0 and RS1:4).

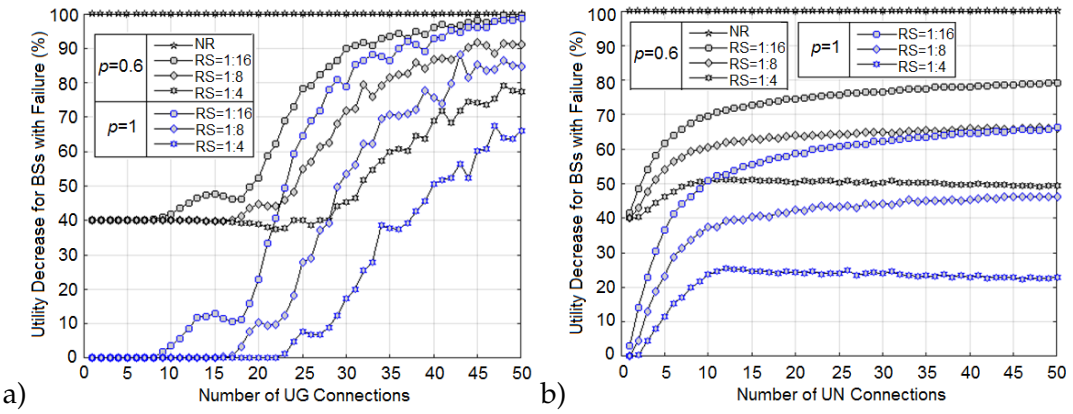

FIGURE 5.7: Average Utility Reduction in BSs Affected by Link Failures. (a) UG Connections. (b) UN Connections.

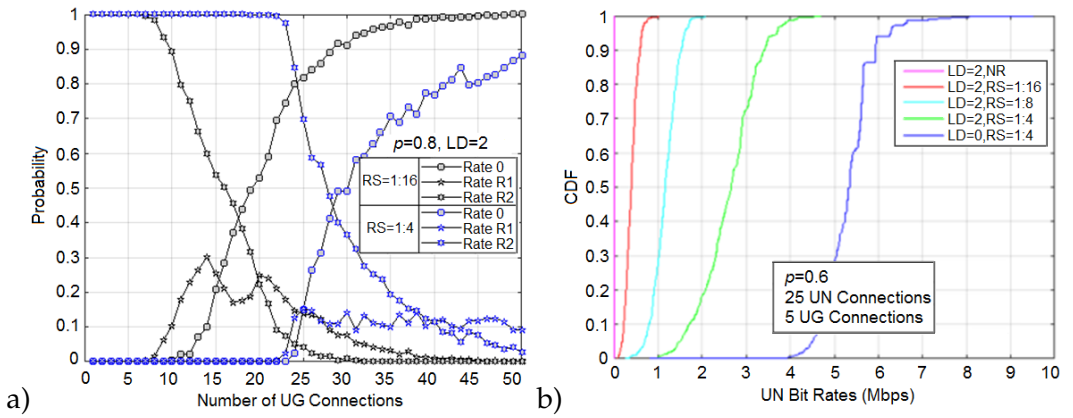

FIGURE 5.8: (a) PDFs of UG Bit Rates for RS=1:16 and RS=1:4. (b) CDF of Bit Rates Assigned to UN Connections.

5.8 Summary

The integration of a satellite component within mobile backhaul networks is regarded as a compelling proposition to increase its availability and resiliency. This chapter has developed and numerically assessed by means of simulation a traffic distribution strategy that exploits a dynamically steerable satellite capacity provisioned for resilience purposes to maximize a network utility function. The analysis has considered different resilience schemes, characterized by the ratio $N : M$ of satellite capacity needed to cope with the failure of N terrestrial links in a group of M BSs. Obtained results show how, under failure conditions, the proposed traffic distribution strategy is able to redistribute the satellite resources so that the utility decrease in the BSs affected by terrestrial link failures is minimized. And remarkably, when all the terrestrial links are fully operational, it has been shown that the proposed strategy does not leave the satellite capacity unused but exploits it in order to improve the overall network utility.

The presented results give us some fundamental guidelines for deriving design and dimensioning principles to consider resilience as a built-in feature in 5G systems. The next chapter considers a deeper analysis of the optimality conditions and algorithms to cope with the presented optimization problem as well as its extension to jointly consider other uses of the 5G backhaul integrated satellite capacity such as fast deployments/special events and the delivery of multicast traffic.

Chapter 6

Formulation and Performance Assessment of a SDN-Based TE Application

6.1 Introduction

In the previous chapter an SDN-based TE application that exploits the allocation criteria capacity that is based on the network utility maximization criteria has been developed and evaluated. Unlike the algorithm presented in the previous chapter, this chapter develops and evaluates new traffic distribution schemes under scenarios where moving cells are integrated to the network and under terrestrial link failures are presented in the network. The proposed TE applications also differ from the previous ones in that they adapt reservation schemes of a certain satellite capacity for the BSs that for some reasons do not have terrestrial link capacity at their disposal. The proposed scheme pursues an optimal allocation of the available satellite and terrestrial capacity so that an overall network utility is maximized under both failure and non-failure conditions in the terrestrial links or mobile cell deployments. Results show how the capacity reserve scheme introduction in the algorithms even increase the network utility, this is also explained in detail in this chapter.

The new SDN-based TE applications also are based in the utility framework model for the analysis of traffic distribution strategies and the capacity dimensioning results presented in the previous sections. It is also considered SDN-based TE applications based in a centralized control for managing dynamically a steerable satellite

capacity that exploits the combination of control features and criteria that include: end-to-end path computation; satellite capacity resource reservations; allocation criteria depending on the traffic nature; admission control and rate control features; and network utility maximization criteria for resilience or emmergency purposes. It is considered unicast and multicast traffic as well as elastic and inelastic traffic.

The rest of the chapter is organized as follows; first, it is provided and descripted the network and traffic model as well as the utility framework utilized. Then, it is presented the integration approach for E2E SDN-Based TE that includes some illustrative TE Workflows to explain the operation of path computation mechanisms. Subsequently, it is presented the formulation of the problem for optimal traffic distribution and it is descripted the heuristic we have relied on for the assessment. Finally, a performance analysis is presented to assess the behavior of the proposed SDN-based TE application under diverse scenarios, including terrestrial link failures in some of the Base Stations (BSs) and deployment of a number of transportable BSs that exclusively rely on the satellite capacity for backhauling, as well as homogeneous and non-homogeneous traffic load situations. The proposed SDN-based TE application is assessed and compared against a traditional overflow solution under different scenarios (different number of BSs with no terrestrial link availability).

6.2 SDN-based TE Application

TE solutions for hybrid terrestrial-backhaul backhaul networks shall be able to use the satellite capacity in the way that best complements the terrestrial capacity in front of the changing conditions of both traffic demand (e.g. increase of traffic demand for an especial event, spatial demand fluctuations over time) and network situation (e.g. backhaul backup for terrestrial link failures, network rapid roll-out, fast response capacity, cells on wheels). Facing these multiple and diverse conditions in a consistent becomes challenging for traffic engineering.

The big advantage of a centralized SDN framework for the realization of TE solutions is that there is a holistic view of the network together with mechanisms to enforce network polices from a single touch point. Based on the SDN architecture for SNS discussed in the chapter 3, Figure 6.1 illustrates the approach adopted for the

realization of the E2E traffic engineering applications spanning the MNO and Satellite Network Operator (SNO) domains when the SNS is used for mobile backhauling services. As depicted in the figure, a centralized software for E2E TE purposes sits on top of the MNO controller analyzing and controlling the E2E network forwarding paths across the satellite and terrestrial backhaul parts. TE applications benefit from a real-time view of the network, which might include insights into various network analytics about link congestion, delay, latency, drops and other performance metrics. For instance, as detailed in chapter 3, OpenFlow supports monitoring statistics that provide visibility into specific application performance. This centralized knowledge is then used to establish and push down rules to the nodes on a per flow basis or guide the network behavior at the desired traffic granularity. Therefore, traffic flowing can be engineered flexibly pursuing one or a combination of goals such as maximizing aggregate network utilization, providing optimized load balancing, minimizing power consumption, and other generic traffic optimization techniques.

It's worth noting that per flow TE state is challenging to achieve with traditional MPLS/TE mechanisms. The main reason is that MPLS/TE is a distributed architecture in which there is no real-time global view of the end-to-end network path. This may induce incorrect traffic engineering decisions, lack of predictability and deterministic scheduling of the label switched paths. While there are tools available that work in conjunction with MPLS TE to create a holistic view, they are usually expensive and do not offer a "real-time" picture [99]. They often create an offline topology and, indeed, they also don't change the fact that MPLS is a distributed architecture. There are possibilities with NETCONF and MPLS-TP but operationally cause problems and don't alleviate the distributed signaling protocols.

SDN-based TE applications have been already proposed mostly for data center or enterprise network scenarios, including ElasticTree [100], Hedera [101], OpenFlow-based server load balancing [102], Plug-n-Serve [103], QNOX [104], QoS for SDN [105], QoS framework [106] and many others (see [107] for a detailed survey). Among the most recent proposals, we find optimization of rules placement [108], the use of MAC as a universal label for efficient routing in data centres [109], among other techniques for flow management, fault tolerance, topology update, and traffic characterization [110].

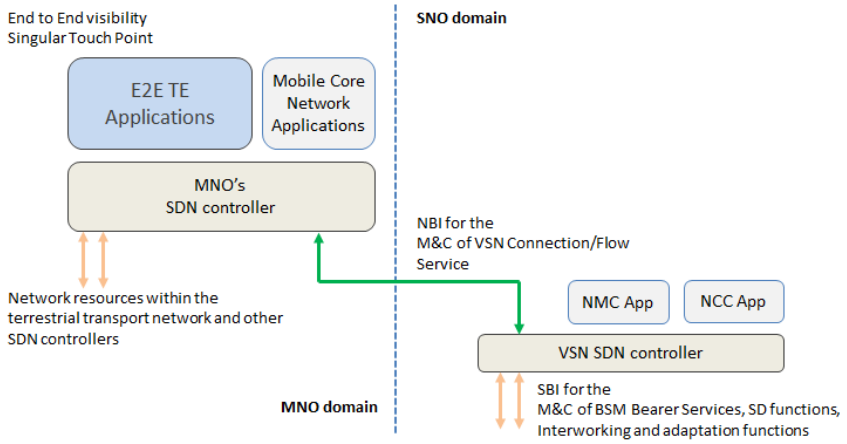

FIGURE 6.1: E2E Traffic Engineering Applications Spanning the MNO and SNO Domains.

In the context of mobile backhauling scenarios, the proposed approach can be used to develop TE applications that exploit the following sort of features and criteria:

- E2E path computation with selection of the terrestrial or satellite link for backhauling, considering multiple optional link utilization rates and flow characteristics more comprehensively.
- Resource reservations mechanisms to protect or give preferential treatment to applications / users / locations such as e.g. BSs with no or limited terrestrial link backhaul capacity that are fully reliant on the use of a pool of shared satellite capacity.
- Different allocation criteria depending on the traffic nature (e.g. guaranteed bit rate or not, unicast or multicast).
- Admission Control and Rate control to face overload and guarantee resources and minimum (committed) transmission rates per flow and group of flows.
- Use of network utility maximization criteria, where the adequacy of handling specific flows over the terrestrial or satellite component, as well the effect of allocating more or less data rate, can be accounted.
- Exploit Bandwidth on Demand (BoD) features if provided by the underlying transport infrastructure.

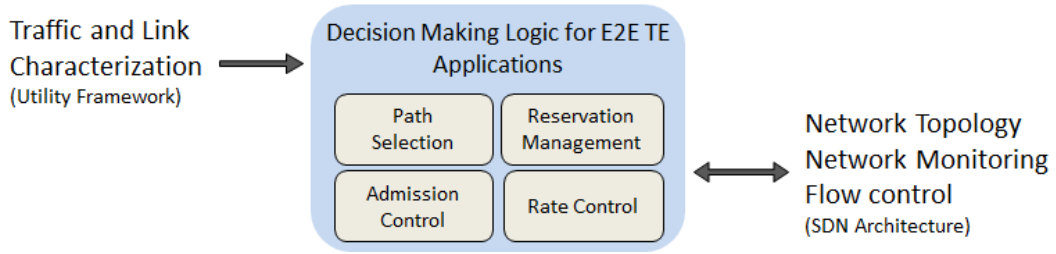

FIGURE 6.2: Components of the E2E Traffic Engineering Application.

- Control the activation/deactivation of networking functions for traffic optimization (e.g. trigger the deployment of Satellite Network Functions (SNFs) for compression, TCP optimization, etc.).

In this context, the TE solution formulated and assessed in this section combines the following control features as part of its decision-making logic: path selection, admission control, rate control and reservation management, as illustrated in Figure 6.2. On this basis, next subsection details the specific traffic and link characterization approach that has been established for the specification of the TE decision-making logic. Then, in the subsequent subsection, the optimization problems and algorithms designed behind the TE decision-making components are described.

6.3 Traffic and Link Characterization for TE

Similarly to network model of previous chapter (See section 5.3), we also consider a cellular network with M Fixed Base Stations (FBS) deployed across a large geographical area where the transport connectivity between the M BSs and the mobile core network is assumed to be delivered through a hybrid satellite-terrestrial backhaul network. However, for this analysis we also consider a number of N BSs, referred to as Transportable BSs (TBS), used for temporary deployments/ fast network roll-out that only rely on the use of the satellite backhaul links.

The traffic model again considers mixes of elastic and inelastic traffic, however, to this analysis we also consider the utilization of multicast GBR services (called MG services in the following). The consideration of MG services in the analysis allows us to exploit the intrinsic broadcast/multicast transmission capacity of the satellite component, assessing its impact on the network in terms of QoS. Unlike unicast

services, one particular MG session can consist of multiple MG flows simultaneously forwarded to multiple BSs simultaneously [80].

Then, the maximum terrestrial and satellite link capacity at the m-th fixed base station (FBS_m) is denoted, respectively, as $C_{FBS-T,m}$ and $C_{FBS-S,m}$, while the satellite link capacity at TBS_m is referred to as $C_{TBS-S,m}$. Accordingly, the aggregate vectors of the such backhaul link capacities are denoted as $\mathbf{c}_{FBS-T} = \{C_{FBS-T,1}, ..., C_{FBS-T,M}\}$, $\mathbf{c}_{FBS-S} = \{C_{FBS-S,1}, ..., C_{FBS-S,M}\}$ and $\mathbf{c}_{TBS-S} = \{C_{TBS-S,1}, ..., C_{TBS-S,N}\}$. In addition, it is assumed that maximum aggregate satellite capacity in use at a given time across all BSs cannot exceed C_S which is defined as $C_S = C_{FBS-S,1} + ... + C_{FBS-S,M} + C_{TBS-S,1} + ... + C_{TBS-S,N}$. Moreover, the terrestrial link availability is captured through a binary vector $\mathbf{a} = \{a_1, ..., a_M\}$ where $a_m = 1$ stands for the terrestrial link at BS_m being operational and $a_m = 0$ represents a Link Down (LD) situation.

Then, let's define $I_S(m)$ as the set of service flows types $s \in \{UG, MG, UN\}$ served through the m-th BS, which refers to FBS_m if $1 \leq m \leq M$ or TBS_{m-M} if $M < m \leq M + N$. Then, let $x_{s,m,i}$ and $r_{s,m,i}$ (denoted as matrices X and R respectively) denote respectively, the link selection (if the flow is served through the satellite backhaul link $[x = 0]$ or through the terrestrial backhaul link $[x = 1]$) and bit rate allocation variables associated with the i-th flow of s service type at the m-th BS. For multicast traffic, let's also define I_{MG} and $Z(j)$ respectively, as the set of MG sessions and the set of BSs to which the MG flows belonging to the j-th MG session are simultaneously forwarded.

The TE decision-making logic consists of a combination of processes, some executed when there is some trigger (e.g. new flow request) and others executed periodically (e.g. performance metric computation and flow adjustments). Figure 6.3 shows the TE decision making logic to handle a UG flow request. At a request arrival, it is checked if the new UG flow is to be served through a BS with terrestrial and satellite links both operational or only with satellite capacity available (e.g. terrestrial link failure situation, transportable BS with no terrestrial backhaul). In the former case, the TE algorithm continues by checking whether there is sufficient capacity across any of the paths to serve the new flow without compromising the quality of the already established GBR flows (UG and MG active flows). This is achieved

by establishing a *GBR Admission Load Threshold* to limit the maximum capacity occupation of a link allowed for use of GBR traffic. This parameter is used by Admission Control #1 and #2 logic (detailed in Table 6.1). If there is sufficient backhaul capacity across both satellite and terrestrial links, the flowchart continues by computing the achievable global network utility (i.e. aggregate of the utility of established flows plus the utility of the new flow) for each of the two candidate paths, selecting the one leading to the higher utility increase. Note that the utility computation is not conducted when there is only a candidate option or when none of them is available, leading the latter case to the rejection of the UG flow request. For the admitted UG flow requests, the GBR and Maximum Bit Rate (MBR) are both set to the rate that gives the maximum utility for UG services.

6.4 TE Decision-Making Logic

The TE decision-making logic consists of a combination of processes, some executed when there is some trigger (e.g. new flow request) and others executed periodically (e.g. performance metric computation and flow adjustments). Figure 6.3 shows the TE decision making logic to handle a UG flow request. At a request arrival, it is checked if the new UG flow is to be served through a BS with terrestrial and satellite links both operational or only with satellite capacity available (e.g. terrestrial link failure situation, transportable BS with no terrestrial backhaul). In the former case, the TE algorithm continues by checking whether there is sufficient capacity across any of the paths to serve the new flow without compromising the quality of the already established GBR flows (UG and MG active flows). This is achieved by establishing a *GBR Admission Load Threshold* to limit the maximum capacity occupation of a link allowed for use of GBR traffic. This parameter is used by Admission Control 1 and 2 logic (detailed in Table 6.1). If there is sufficient backhaul capacity across both satellite and terrestrial links, the flowchart continues by computing the achievable global network utility (i.e. aggregate of the utility of established flows plus the utility of the new flow) for each of the two candidate paths, selecting the one leading to the higher utility increase. Note that the utility computation is not conducted when there is only a candidate option or when none of them is available, leading the latter

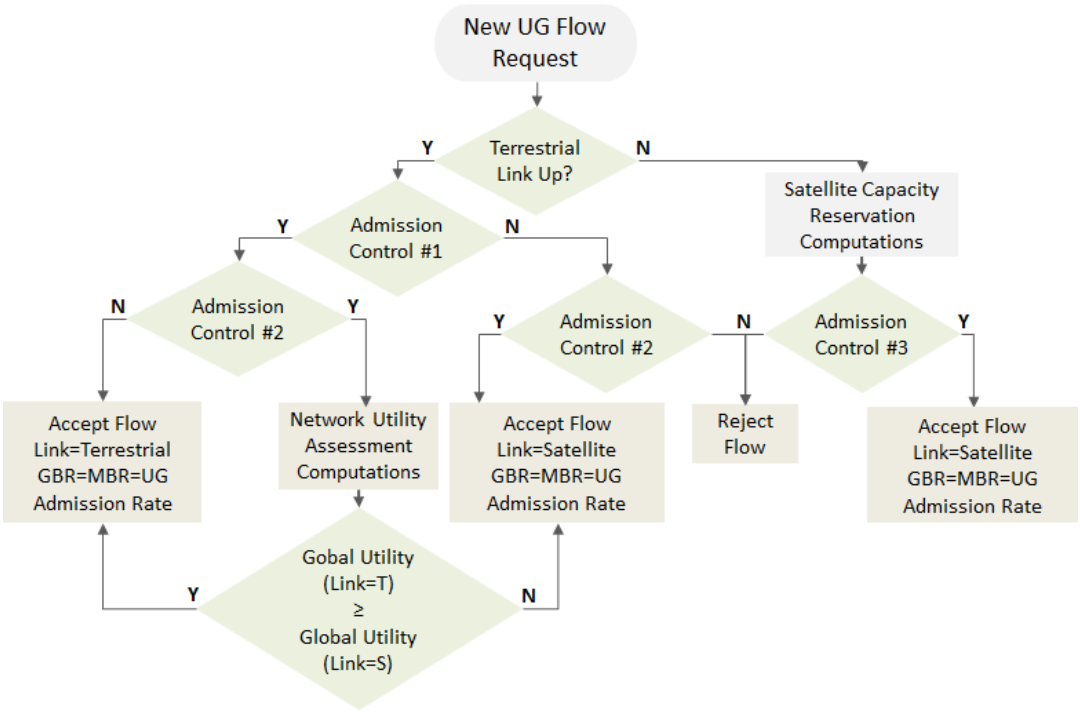

FIGURE 6.3: TE Decision-Making Logic to Handle New UG Flow Arrivals.

case to the rejection of the UG flow request. For the admitted UG flow requests, the GBR and Maximum Bit Rate (MBR) are both set to the rate that gives the maximum utility for UG services.

As previously noted, the flowchart in Figure 6.3 also captures the case where the UG flow is to be served through a BS where the terrestrial link is not available. In this regard, as seen on the right side of the flowchart, a resource reservation management mechanism is introduced in the decision-making process. This mechanism is used to enforce a preferential treatment for the use of the shared satellite capacity to the BSs without an operational terrestrial link. Therefore, at a new UG flow request arrival, now the TE logic first goes through Admission control #3 (detailed in Table 6.1) that takes into account the amount of reserved satellite capacity that is dynamically adjusted over time for the serving BS. The computations needed to manage such satellite capacity reservations are detailed later on in this section.

The TE decision making logic to handle UN flow requests is depicted in Figure 6.4. Similar to the treatment of UG flows, the TE algorithm first checks if the new

TABLE 6.1: Admisssion Control Computations.

Admission Control	**Description**
Admission Control 1	(GBR Terrestrial Load at BS + UG Admission Rate)< (GBR Admission Load Threshold · Terrestrial Link Capacity at BS)
Admission Conmtrol 2	(GBR Satellite Load at BS + UG Admission Rate) < (GBR Admission Load Threshold · Satellite Link Capacity at BS) AND (Global GBR Satellite Load + UG Admission Rate) < (GBR Admission Load Threshold · (Satellite System Capacity – Satellite Reserved Capacity)
Admission Control 3	GBR Satellite Load at BS + UG Admission Rate) < (GBR Admission Load Threshold Satellite Reserved Capacity at BS)

Where:
UG Admission Rate: Rate that is considered in the admission process. It is selected from the rates specified for the definition of the utility functions.
Satellite system capacity (C_S): Total amount of satellite capacity shared by a group of BSs.
Satellite reserved capacity (Cr): Satellite capacity reserved for preferential reserved for preferential use of a given BS.
GBR Admission Load Threshold: Maximum percentage of the available (satellite, terrestrial, reserved) capacity that can be used to serve GBR traffic.

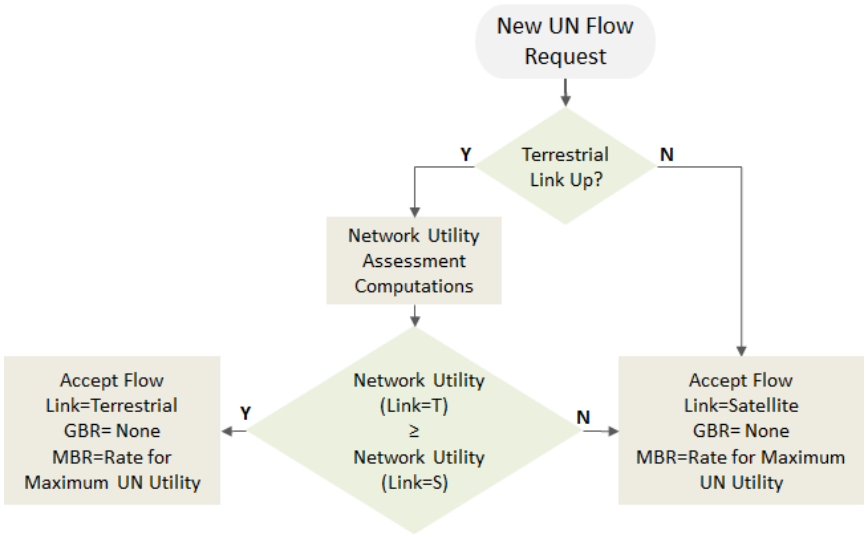

FIGURE 6.4: TE Decision-Making Logic to Handle New UN Flow Arrivals.

UN flow is to be served through a BS with terrestrial and satellite links both operational or only with satellite capacity available. In the former case, the next step is to compute the overall utility increase that would be achieved if the flow is enforced through the terrestrial or satellite links, selecting the option that turns into the higher network utility increase. In the latter case, shown on the right side of Figure 6.4, the flow is always enforced through the satellite connection and the reservation amount is updated accordingly. Note that, unlike UG flows processing, no admission control is enforced for UN flows because of its elastic traffic nature (i.e. the rates achieved per flow are variable and depend on overall number of flows simultaneously served in the network). Therefore, no GBR rate is established for the admitted flows and the MBR parameter, used for rate control purposes, is set to the rate that achieves the maximum utility for UN services.

Even though network utility maximization is sought after each flow arrival, traffic variations (e.g. termination of established flows) and changes in capacity conditions (e.g. changes in reservations, terrestrial link failures) might turn into situations that the achieved network utility is not optimal. To face this situation, a mechanism to re-assess the network utility of the established flows and, if necessary, carry out any re-allocations is considered. This process is illustrated in Figure 6.5. As seen in the figure, network utility re-assessment and reallocation is triggered periodically as

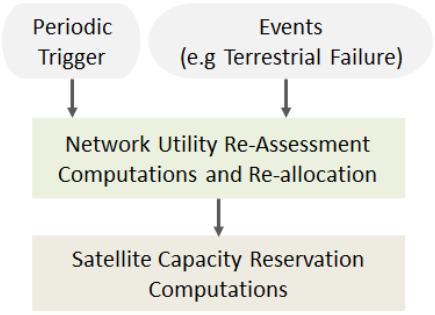

FIGURE 6.5: Logic for Continuous Network Utility Re-assessment, Reallocation and Reservation Update.

well as due to the occurrence of specific events such as a change in the amount of capacity in a network link. Figure 6.5 also shows that, after the execution of the network utility re-assessment and reallocation process, capacity reservations are also revisited to account for any changes enforced to the on-going flows.

Capacity Reservation Computations

The reservation management mechanism aims to ensure that some amount of satellite capacity remains available for the BSs that do not count with terrestrial capacity. Indeed, considering that one of the conditions that lead to global utility maximization is a fair distribution of the rates delivered to UN flows. This reservation mechanism helps in achieving fairness in terms of the overall capacity distribution among BSs (i.e. BS without terrestrial capacity will get a higher share of the satellite capacity). To that end, the satellite reserved capacity (C_r) variable is introduced. This parameter is initialized with a default reservation value and periodically updated over time based on the evolution of the traffic load served through the corresponding BS (details are given in Table 6.2). In particular, C_r is computed to account for the UG traffic load supported at the BS plus an additional capacity for UN traffic that would allow to deliver an average bit rate as that achieved across the whole network for UN flows. The value of C_r is constrained by the terrestrial link capacity at the BS, the *Maximum Capacity Reservation* per BS and the *Maximum Capacity Reservation* applicable to the total satellite reserved capacity. The remaining satellite system resource available for BSs with terrestrial capacity are defined as satellite non-reserved Capacity (C_{nr}).

TABLE 6.2: Satellite Reservation Computation

Reservation Control Parameters	**Computation**
Satellite reserved capacity at $BS_m(Cr_m)$	(Cr_m =UG Satellite Load at BS_m + UN Satellite Flows at BS_m· UN Global Average Flow Rate
Satellite non-reserved Capacity (Cnr)	Cnr=Satellite system capacity (C_S) - $\sum_{BS_m} Cr_m$
Constraints: Total satellite reserved capacity $\leq$ Maximum Capacity Reservation Satellite reserved capacity at $BS_m \leq$ Maximum Capacity Reservation per BS Satellite reserved capacity at $BS_m \leq$ Terrestrial Link Capacity at BS_m	

Utility Maximization Computations

For the computation of the maximum achievable utility, each time a new flow request is received and as part of the periodic network utility re-assessment and reallocation process, the following optimization problem is solved.

On this basis, the optimization problem can be formulated as basis:

Given: FBS, TBS (Set of Fixed and Transportable BSs)
$C_{FBS-T}, C_{FBS-S}, C_{TBS-S}, C_S$, **a** (Link capacity vectors, satellite system capacity and availability vector, respectively)
$I_s(m), I_{MG}, Z(j)$ (Set of service flows types, the set of MG sessions and the set of BSs to which the MG flows respectively)

Find: $\{x_{s,m,i}\}, \{r_{s,m,i}\}$(Link Selection and Bit Rate Variables)

Max:

$$\sum_{s\in\{UG,MG,UN\}} \sum_{m=1}^{N+M} \sum_{i\in I_s(m)} U^S\left(r_{s,m,i}, x_{s,m,i}\right) \tag{6.1}$$

s.t.

$$\sum_{s\in\{UG,MG,UN\}} \sum_{i\in I_s(m)} r_{s,m,i}\cdot x_{s,m,i} \leq a_m \cdot C_{FBS-T,m} \qquad m=1,...,M \tag{6.2}$$

$$\sum_{s\in\{UG,MG,UN\}} \sum_{i\in I_s(m)} r_{s,m,i}\cdot (1-x_{s,m,i}) \leq C_{FBS-S,m} \qquad m=1,...,M \tag{6.3}$$

$$\sum_{s\in\{UG,MG,UN\}} \sum_{i\in I_s(m)} r_{s,m,i} \leq C_{TBS-S,m-M} \qquad m=M+1,...,M+N \tag{6.4}$$

$$x_{s,m,i}=0 \quad \forall s,m,i : i\in I_S(m) \qquad m=M+1,...,M+N \tag{6.5}$$

$$\sum_{s\in\{UG,UN\}}\sum_{m=1}^{N+M}\sum_{i\in I_s(m)} r_{s,m,i,}\cdot(1-x_{s,m,i}) \quad + \\ \sum_{j\in I_{MG}}\max_{\forall i,m|i\in I_{MG}(m)\cap BS_m\in Z(j)}(r_{MG,m,i}\cdot(1-x_{MG,m,i}))\leq C_S \tag{6.6}$$

$$x_{s,m,i}\in\{0,1\} \quad \forall s,m,i \quad ,m=1,...,M \qquad r_{s,m,i}\in\left\{\Re^+\right\},\forall s,m,i \tag{6.7}$$

where Eq. 6.1 is the objective function defined as the sum of all service flows utilities, Eq. 6.2 - 6.3 represent, respectively, the terrestrial and satellite link capacity constraints at each FBS_m, Eq. 6.4 represents the satellite link capacity constraint at each TBS_m while Eq. 6.5 indicates that the routing variable for all service flows belonging to the N TBSs will be pre-set to be handled via satellite since these BSs do not have terrestrial links availability. The Eq. 6.6 indicates the total satellite capacity constraint, denoting for MG traffic, as the satellite system capacity consumed by the j-th MG session is equal to the highest bit rate allocated to any of the constituent service flows that are handled via satellite. Finally, Eq. 6.7 indicates the possible variable values.

The above optimization problem is a non-linear optimization problem (NLP) with an objective function that combines concave and discontinuous utility functions and constraint functions that contain products of the decision variables. Some authors dealing with rate allocation schemes for networks sharing these both services types agree on the difficulty of solving this type of network utility maximization problems [111]-[114]. Indeed, instead of dealing directly with the complete problem, they propose alternative rate allocation schemes, separating and independently managing the resource allocation for inelastic and elastic traffic, limiting the inelastic traffic allocation to a pre-set link utilization percentage and solving the network utility maximization problem for elastic traffic considering the remaining links available resources.

Based on above considerations, we have solved the network utility maximization problem by splitting it in three separate sub-problems. In particular, the problem is solved sequentially first considering only UG services (sub-problem # 1), then adding MG services (sub-problem # 2) and finally adding the UN services (sub-problem # 3). Some details on the algorithms used to solve each of the sub-problems are provided in the following.

Sub-problem # 1 and # 2 can be indeed formulated as the general problem in Eq. 6.1-6.7 with $s \in \{UG\}$ and $s \in \{MG\}$, respectively. Each of these sub-problems is a discrete optimization problem (note that utility functions for both unicast and multicast GBR traffic are step functions) that is solved by means of a heuristic. The proposed heuristic is based on a greedy algorithm, which, flow by flow, decides on the link and bit rate that yields the highest utility. In this way, as long as capacity constraints are met, the algorithm always moves towards getting the maximum utility. When capacity constraints are found and a flow cannot be accepted with its maximum utility, the algorithm revisits some of previous decisions in order to see if some modifications are possible (e.g. rate reduction or link change) as long as these modifications result into higher network utility. The algorithm is detailed in the Algorithm 1. For the allocation and re-allocation decisions, the algorithm uses a set of pre-computed utility increments ΔU_n that account for the potential combinations of link and rate selection of the current and up to q previous allocated flows. In the case of sub-problem for UG services, these increments are computed as follows: $\Delta U_1 \leftarrow U^{UG}(R_2^{UG},1)$, $\Delta U_2 \leftarrow U^{UG}(R_2^{UG},0)$, $\Delta U_3 \leftarrow U^{UG}(R_1^{UG},1)$ and $\Delta U_4 \leftarrow U^{UG}(R_1^{UG},0)$, for allocations that do not impact on previous flows and $\Delta U_5 \leftarrow 2U^{UG}(R_1^{UG},1)$-$U^{UG}(R_2,1)$, $\Delta U_6 \leftarrow 2U^{UG}(R_1^{UG},0)$- $U^{UG}(R_2^{UG},0)$, ... , $\Delta U_{n-1} \leftarrow (q+1)U^{UG}(R_1^{UG},1)$-$qU^{UG}(R_2^{UG},1)$, $\Delta U_n \leftarrow (q+1)U^{UG}(R_1^{UG},0)$-$qU^{UG}(R_2^{UG},0)$, for allocations that force a rate reduction in previously allocated flows. Also is worth noting that the computation of the available capacity considers the impact of the satellite reservation parameters in the BSs with no terrestrial capacity available so that the following settings $C_{TBS-S,m-M} = Cr_m$ and $C_s = Cr_m$ are considered when conducting the resource allocation for the flows served by these BSs.

Procedure: Resource Allocation (service type (s), $S_{BS} \in \{FBS, TBS\}$***)***

$\Delta U = \{\Delta U_1, \Delta U_2, ..., \Delta U_n\}$, Set of utility increments different potential allocations

$r_{s,m,i} \leftarrow 0$ for all *m,i*

for $BS_m \in S_{BS}$ **do**

  $AvailableCapacity = TRUE$

  for $i \in I_s(m)$ *and while Available Capacity=TRUE* **do**

    Select $(r_{s,m,i}, x_{s,m,i})$ and the applicable rate deruction of previous allocated flows yielding the highes $\Delta U_n \in \Delta U$ that satisfies Eq.6.2-6.7,

    if $(r_{s,m,i} = 0)$, $AvailableCapacity = FALSE$

  end

end

Algorithm 3: Resource Allocation Algorithm for UG Service Flows.

Sub-problem #3 for utility maximization of UN services can also formulated from Eq. 6.1 – 6.7 with $s \in \{UN\}$ and the following new link capacities.

$$C'_{FBS-T,m} = C_{FBS-T,m} - \sum_{s\in\{UG,MG\}} \sum_{i\in I_s(m)} r_{s,m,i} \tag{6.8}$$

$$C'_{FBS-S,m} = C_{FBS-S,m} - \sum_{s\in\{UG,MG\}} \sum_{i\in I_s(m)} r_{s,m,i} \tag{6.9}$$

$$C'_{TBS-S,m-M} = C_{TBS-S,m-M} - \sum_{s\in\{UG,MG\}} \sum_{i\in I_s(m)} r_{s,m,i} \tag{6.10}$$

$$C'_S = C_S - \sum_{m=1}^{N+M} \sum_{i\in I_{UG}(m)} r_{UG,m,i} \quad - \quad \sum_{j\in I_{MG}} \max_{\forall i,m | i\in I_{MG}(m)\cap BS_m\in Z(j)} r_{MG,m,i} \cdot (1 - x_{MG,m,i}) \tag{6.11}$$

In this case, sub-problem #3 turns to be an optimization problem which combines a concave utility function with a discrete link selection variable, with constraint functions that contain products of both concave and discrete optimization variables. Despite the problem shows complexity regarding the application of a direct method for its solution, the consideration of a single type of traffic $s \in \{UN\}$ facilitates its analysis and the introduction of a heuristic. Indeed, in [95] it is demonstrated that a proportional fair resource allocation ensures utility max-min fairness for all users

sharing a single path in the network, i.e. the utility that should be shared fairly among users. Then, a max-min rate allocation is equivalent to the maximization of the utility at each link for UN traffic. This same concept can be extended to the global utility of the network when there is no distinction between using satellite or terrestrial capacity, which in our case is strictly valid for factor $p^{UN} = 1$. Since in the network model a service flow can be handled only by a single link, and in turn each BS has at most two possible links, for $p^{UN} = 1$, a heuristic can find the utility maximization applying the above mentioned rule, selecting the link to each UN flow in such a way that all service flows have as much as possible the same bit rate utilizing the full links capacities. To do this, the algorithm solve the number of UN flows handled by terrestrial or satellite links at each BS_m(K_m^T and Km^S respectively) that theoretically would give the same bit rate to all UN connections by solving the Eq. 6.12 - 6.13:

$$r_{UN,m,i} = \frac{C'_{FBS-T,m}}{K_m^T} = \frac{C'_{FBS-S,m}}{K_m^S}, \quad \forall i : i \in I_{UN}(m), \quad m = 1, ..., M \tag{6.12}$$

$$r_{UN,m,i} = \frac{C'_{TBS-S,m-M}}{K_m}, \quad \forall i : i \in I_{UN}(m), \quad m = M+1, ..., M+N \tag{6.13}$$

where

$$r_{UN,1,i} = r_{UN,2,i} = \cdots = r_{UN,M+N,i}, \qquad \forall i \tag{6.14}$$

s.t. Eq. 6.2 and 6.7.

where K_m is defined as the total UN service flows at the BS_m. For $p^{UN} < 1$, besides solving Eq. 6.12 - 6.14 and fixing the satellite links utilized satellite capacities $C'_{TBS-S,m-M}$ and $C'_{FBS-S,m}$, the algorithm reselects the right links $x_{UN,m,i} i \in I_{UN}(m)$, in order to get K_m^T and K_m^S that maximize the sum of UN service flows utilities at each BS_m. The procedure is detailed in Algo. As in the cases for GBR services, satellite reservation parameters in the BSs with no terrestrial capacity available are accounted by setting $c_{TBS-S,m-M} = Cr_m$ and $C_S = Cr_m$ when conducting the resource allocation for the flows served by these BSs.

Procedure: Resource Allocation (UN flows, $S_{BS} \in \{FBS, TBS\}$)

$r_{UN,m,i} \leftarrow 0$ for all m, i

for $BS_m \in S_{BS}$ **do**

  K_m^T,K_m^S,$C'_{TBS-S,m}$,$C'_{FBS-S,m} \leftarrow$ Solve Eq. 6.12 - 6.14;

end

for $BS_m \in S_{BS}$ **do**

  for $i \in \{I_{UN}(m)\}$ *and while Available Capacity=TRUE* **do**

    Given $C'_{TBS-S,m}$,$C'_{FBS-S,m}$, find (K_m^T, K_m^S) that yields the highest

    $\Delta U_n \in \Delta U$ that satisfies Eq.6.2 - 6.7, this is,

    $\{x_{UN,m,i}\}$,$\{r_{UN,m,i}\}$,K_m^T,$K_m^S \leftarrow argmaxU(x_{UN,m,i}, r_{UN,m,i})$;

  end

end

Algorithm 4: Resource Allocation Algorithm for UN Service Flows.

Reference TE Strategy for Comparison Purposes

For comparison purposes, the assessment presented in the following section considers also a more conventional overflow strategy that is executed locally at each BS and lacks of any centralized control. A state diagram that describes the operation of the overflow strategy is depicted in Figure 6.6. It is considered that each BS with both terrestrial and satellite capacity can switch between two operational overflow states: OFF and ON. In OFF state, all generated backhaul traffic is handled through the terrestrial link. Otherwise, backhaul traffic generated when the BS is in ON state is always directed through the satellite link. As captured in Figure 6.6, when terrestrial capacity is not available, the operation mode remains in the ON overflow state.

The transition between the OFF and ON state is established based on a two-fold condition (Condition1 in Figure 6.6: the amount of GBR load (UG and MG flows) has started exceeding a given threshold (*Overflow GBR Load Activation Threshold*) or the average rate being delivered to UN flows has fallen below a given threshold (*Overflow UN Rate Activation Threshold*). The change is executed if this condition holds for an overflow decision interval (ΔT). Similarly, the transition from the ON to OFF states (Condition2 in Figure 6.6 is determined by the counterpart two-fold condition:

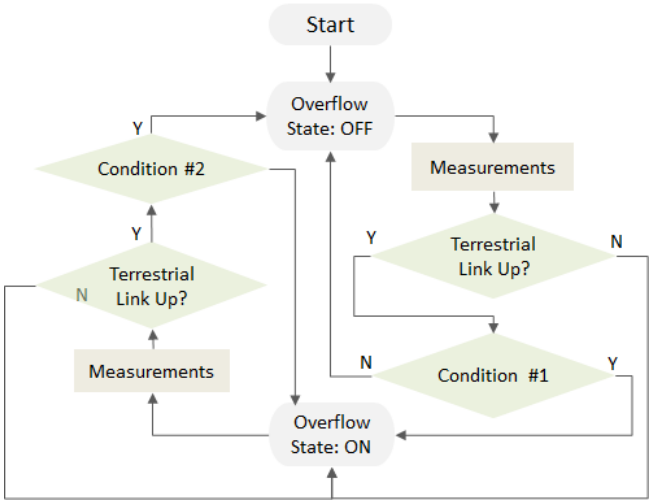

FIGURE 6.6: State Diagram of the Overflow Strategy.

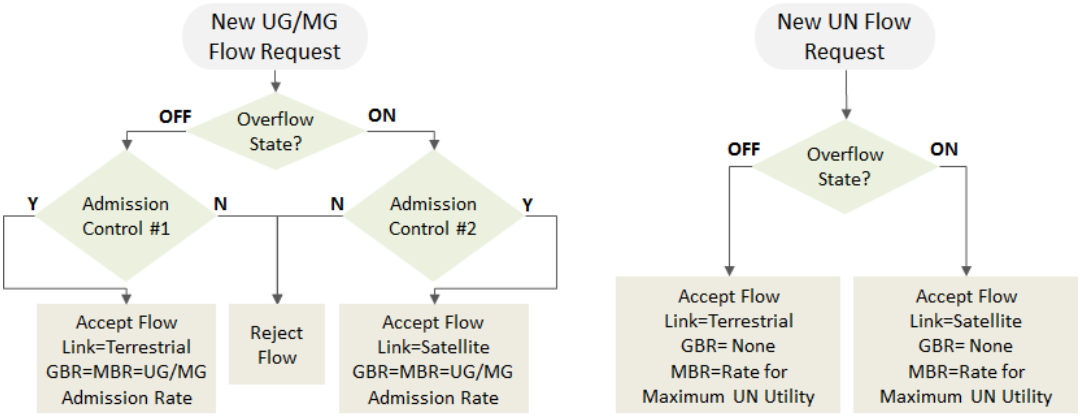

FIGURE 6.7: Logic to Handle Flow Requests Under the Overflow Strategy.

the GBR load has decreased below a given threshold (*Overflow GBR Load Deactivation Threshold*) and the average rate being delivered to UN flows is above a given threshold (*Overflow UN Rate Deactivation Threshold*). Both conditions are detailed in Table 6.3.

Figure 6.7 depicts the flowcharts to handle a UG/MG and UN flow requests under the overflow strategy. Admission control applied in the case of GBR traffic follows the same principles used for the SDN-based TE application. The corresponding admission control computations for the overflow strategy are detailed in Table 6.4.

The centralized control, be it configuration or policy management or traffic engineering significantly simplifies hard-to-solve problems that benefit from centralized visibility. Equipped with global views of network resources and global control of

TABLE 6.3: Overflow States Switching Conditions and Parameters.

State Condition	Computation
From OFF to ON (Condition # 1)	GBR Load level > Overflow GBR Load Activation Threshold OR Average UN Flow Rate < Overflow UN Rate Activation Threshold within [t,t-ΔT]
From ON to OFF (Condition # 2)	GBR Load level > Overflow Load Deactivation Threshold AND Average UN Flow Rate < Overflow UN Rate Dectivation Threshold within [t,t-ΔT]
Parameters: Overflow GBR Load Activation Threshold Overflow GBR Load Dectivation Threshold Overflow UN Rate Activation Threshold Overflow UN Rate Dectivation Threshold Δ T=Overflow decision internal (s)	

TABLE 6.4: Admission Control Computations for the Overflow Strategy

State Condition	Computation
Admission Control # 1	(GBR Terrestrial Load at BS + UG Admission Rate) < (GBR Admission Load Threshold $TerrestrialLinkCapacityatBS$)
Admission Control # 2	(GBR Satellite Load at BS + UG Admission Rate) < (GBR Admission Load Threshold · Satellite Link Capacity at BS) AND (Global GBR Satellite Load + UG Admission Rate) < (GBR Admission Load Threshold · Satellite System Capacity)
Where: UG Admission Rate: Rate that is considered in the admission process. It is selected from the rates specified for the definition of the utility functions. Satellite system capacity (C_S): Total amount os satellite capacity shared by a group of BSs. Satellite reserved capacity (C_r): Satellite capacity reserved for preferential use of a given BS. GBR Admission Load Threshold: Maximum percentage of the available (satellite, terrestrial, reserved) capacity that can be used to serve GBR traffic.	

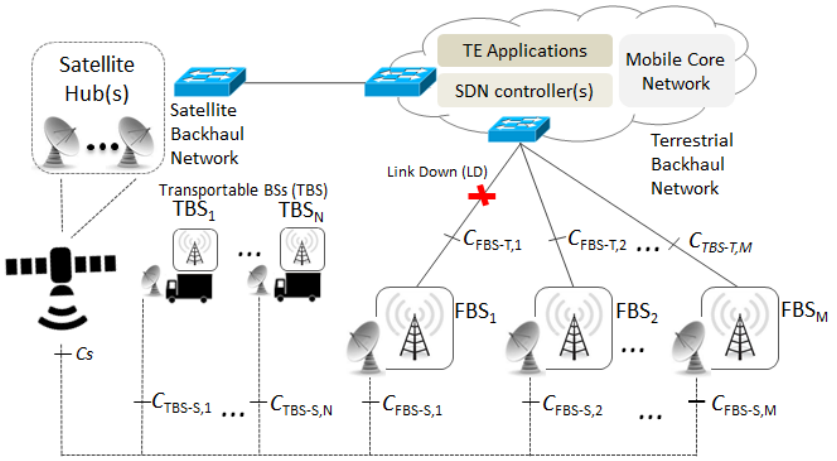

FIGURE 6.8: Evaluated Scenario.

the state of those resources, SDN controllers can make efficient resource allocation decisions in response to changing service demands, and to changing management, policy, and other inputs.

6.5 Numerical Assessment

6.5.1 Simulation Method and Scenario Settings

The simulation scenario considers a set of Base Stations (BSs) with terrestrial and/or satellite backhaul capacity that serve a mix of UG, MG and UN flows. As illustrated in Figure 6.8, some BSs are deployed at fixed locations, referred to as FBSs, with both satellite and terrestrial backhaul links and others, referred to as Transportable BSs (TBS) are used for temporary deployments/ fast network roll-out and only rely on the use of the satellite backhaul links. Table 6-5 provides the range of values considered for the general network deployment settings, the traffic load generation and the configuration of the Overflow and SDN-based TE application in the numerical assessment as well as the values used as default values unless stated otherwise.

In this section, the behavior of the proposed SDN-based TE application is assessed by means of numerical simulations under diverse scenarios, including homogeneous and non-homogeneous load situations, terrestrial link failures in some of the BSs and deployment of a number of transportable BSs that exclusively rely on the satellite capacity for backhauling.

Table 6.5 also provides the range of values considered for admission control and overflow activation/deactivation as *UG admission load threshold* which determines the maximum accepted occupation of GBR traffic, the *Overflow UG Load Activation Threshold* which determines the maximum terrestrial link occupation (of GBR) that activates the overflow of such kind of traffic, the *Overflow N-GBR Rate Activation/Deactivation Thresholds* which determines the minimum/maximum N-GBR bit rates that activates or deactivates respectively the N-GBR traffic overflow and finally the *Overflow Hysteresis* defined as the time after the GBR admission load threshold is exceeded to activate the GBR overflow.

Table 6.5 provides the range of values considered for the traffic load generation and network model parameters in the numerical assessment. With regard to the capacity of the terrestrial links, the considered setting (131 Mbps) is based on the dimensioning analysis presented in chapter 4 to cope with the 90-th percentile of the traffic demand when considering a realistic traffic model that exhibits a log-normal distribution with an average load of 100 Mbps per BS. This value is then considered to establish the range of values for the maximum aggregate satellite capacity (C_S). On the other hand, the maximum satellite link capacity per BS is also set to 131 Mbps in line with the terrestrial capacity and considering that today's top-of-the-line satellite modems based on DVB-S2X can afford this capacity [94]. Both GBR and N-GBR traffic flows are modeled by a Poisson arrivals and exponential session duration disribution and the numerical assessment considers an execution interval of 1000 sec.

All traffic flows are modeled by a Poisson arrivals and exponential session duration distribution. Numerical results have been obtained by running 50 times an event-driven simulation, each representing an execution interval of 1000 sec. The network simulation as well as traffic distribution strategies are developed in Matlab.

For comparison purposes, a traditional overflow strategy is considered. The strategy seeks to emulate a backhaul network with the capacity of overflow activation based only on the occupancy and bit rates measurements achieved locally (in the same BS), this is, a distributed decision making without evaluating the global network measurements (satellite and terrestrial).

TABLE 6.5: Scenario Settings for the Numerical Assessment.

Parameter	Values
General Network Parameters	
Number of BSs sharing the satellite capacity $(N + M)$	16
Number of BSs with no terrestrial link availability	0-4
Backhaul Capacity	
Terrestrial link capacity per $BS_j(C_{FBS-T,m})$	131 Mbps
Maximum Satellite capacity per $BS_j(C_{FBS-S,m}, or, C_{TBS-S,m})$	210 Mbps
Satellite System capacity per C_S (% of terrestrial capacity)	0-10-20% (209.5-419.2 Mbps)
Service flows characterization	
Standard quality UG bit rate R_1^{UG}	3 Mbps[a]
High quality UG bit rate R_2^{UG}	6 Mbps[b]
MG bit rate R_1^{MG}	6 Mbps
UN bit rate for maximum utility R_1^{UN}	13 Mbps
Satellite utility reduction factors (p^{UG}, p^{MG}, p^{UN})	0.6, 0.8, 1.0
UG rate selection	Only High Quality
GBR mean flow duration (sec)	30
UN mean flow duration (sec)	20
Overflow strategy parameters	
GBR Admission Load Threshold	90%
Overflow GBR Load Activation Threshold	80%
Overflow GBR Load Deactivation Threshold	70%
Overflow UN Rate Activation Threshold	40% of R_1^{UN}
Overflow UN Rate Deactivation Threshold	60% of R_1^{UN}
Overflow decision interval ΔT (sec)	5
SDN-based TE application parameters	
GBR Admission Load Threshold	90%
Maximum Capacity Reservation per BS (% of $C_{FBS-T,m}$)	100 %
Initial Capacity Reservation (% of C_S)	20%
Maximum Capacity Reservation (% of C_S)	95%
Re-assessment Update Interval (sec)	1

[a] Typical mobile Video Resolution and Bitrates [85].
[b] The global average for LTE download speed [86].

6.5.2 Illustration of the Operation of the Simulator

Before delving into the key performance indicators achieved per service, Figure 6.9 to 6.12 show some illustrative examples of the dynamic behavior of the system. In particular, the time evolution of the terrestrial and satellite link loads over an execution interval of 1000 seconds in a given BS are depicted in Figures 6.9 and 6.10 for the overflow and SDN-based strategies, respectively. These results are obtained considering a low load condition (low UG load, $\lambda_{UN} = 0.25$) and $C_S = 20\%$ to illustrate the different patterns in the occupation of the satellite link capacity shown by the two strategies: while the overflow only relies on the satellite capacity occasionally, the SDN use it as needed as long as it increases the overall network utility. In the comparison of the satellite occupation, we can notice that the SDN-based TE strategy shows a greater exploitation of the satellite resource. As will be seen in detail later, a better exploitation of satellite capacity as well as a better distribution in terms of UN fairness will allow the SDN-based strategy to achieve better performance in different metrics.

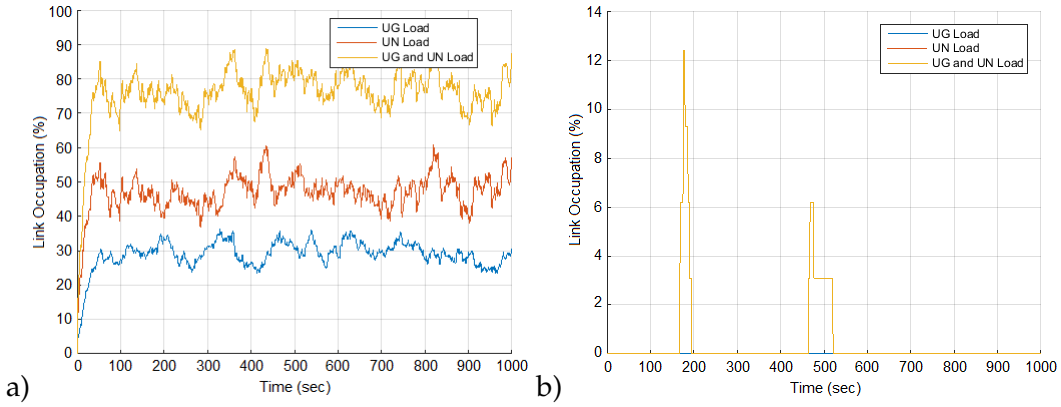

FIGURE 6.9: Time Evolution of the Terrestrial (a) and Satellite (b) Link Capacity Load under the Overflow Strategy.

Figure 6.11 shows an example of the transition pattern between ON/OFF states over time in the overflow strategy under high load conditions. It can be seen that the BS enters into the overflow state multiple times during the execution interval. The two colors shown in the figure are intended to illustrate which condition has triggered the overflow mode: the amount of UG load has exceeded the Overflow GBR Load Activation Threshold or the average rate being delivered to UN flows has

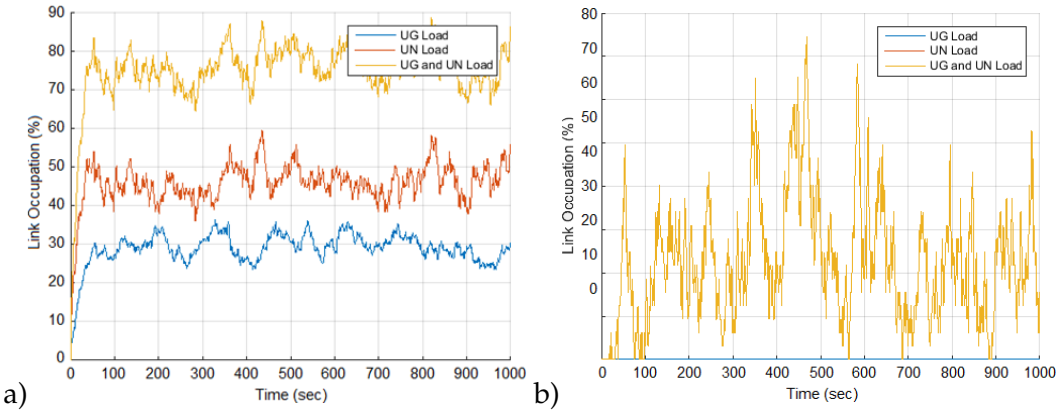

FIGURE 6.10: Time Evolution of the Terrestrial (a) and Satellite (b) Link Capacity Load under the SDN-based TE Strategy.

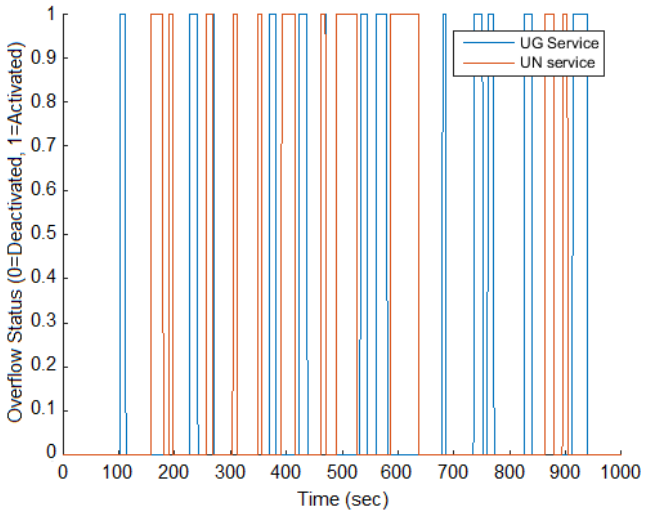

FIGURE 6.11: Transition Pattern of the Overflow on/off Status.

fallen below the Overflow UN Rate Activation Threshold.

Lastly in these series of time evolution illustrative figures, Figure 6.12 shows how the UG rejection events are presented during the execution period for both strategies. Results are obtained in this case considering a medium UG load, $\lambda_{UN} = 1$ and $C_S = 20\%$.

6.5.3 Homogeneous Spatial Traffic Distribution

This first assessment is intended to show the performance of the SDN-based TE application under homogeneous spatial traffic distributions and considering that all BSs have both terrestrial and satellite backhaul capacity. Traffic load for UG services is set to 30% (Low), 60% (Medium) and 90% (High) of the terrestrial link capacity

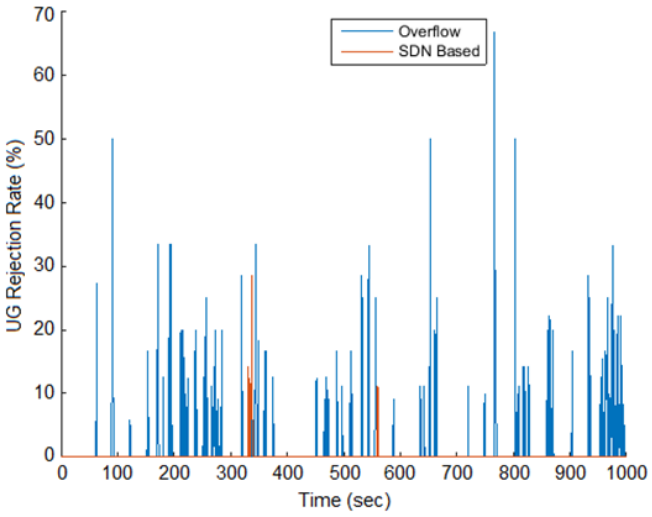

FIGURE 6.12: Time Pattern of the UG Admission Rejections.

in each BS. Considering that UG flows are served with the High quality UG bit rate R_2^{UG} and average session duration is 30s, the corresponding flow arrival rates λ_{UG} for the Low, Medium and High UG load conditions are, respectively, 0.2183, 0.4366 and 0.655 flow/s.

With respect to UN traffic load, the UN service flow arrival rate λ_{UN} is varied between 0.25 and 1.0 at each BS. This results in an average number of active UN flows per BS between 7.5 and 30 per BS considering the average session duration of 20s.

Note that if UN flows could all be served at R_1^{UN}, this would represent an average UG load per BS between 65 and 260 Mbps. No multicast traffic is considered in this first assessment.

Figure 6.13 shows the admission rejection rate experienced by the UG traffic under the SDN-based and Overflow strategies for different amounts of C_S and when considering a satellite utility reduction factor given by $p^{UG} = p^{UN} = 1$. It can be seen how the availability of the satellite capacity leads to a considerable reduction of the rejection rate for UG traffic and how the SDN-based solution clearly outperforms the overflow strategy. For medium UG load, the SDN-based TE application keeps the blocking ratio well below 0.5% with only $C_S = 10\%$ while the overflow strategy is not able to reduce it from 2.0%.

Focusing now on UN service performance indicators, Figure 6.14 shows the mean

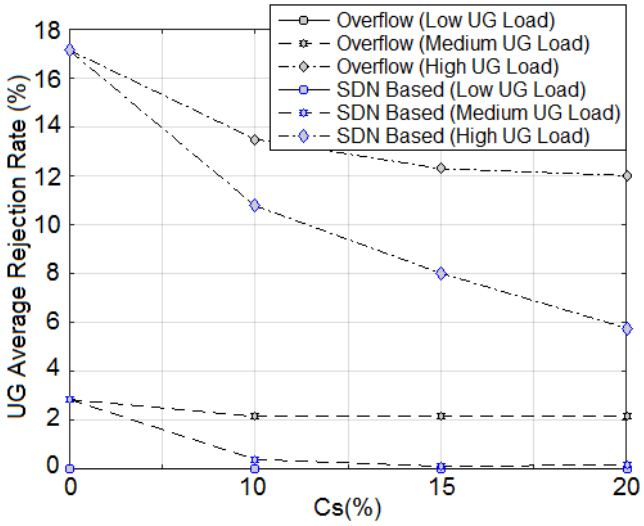

FIGURE 6.13: Admission Rejection Rate for UG Services.

and standard deviation of the data rate delivered per UN flow for different UN loads and considering values of $C_S = 10\%$ and $C_S = 20\%$. Results are obtained for a UG medium traffic load, $p^{UG} = p^{UN} = 1$ and the case with $C_S = 0\%$ is added for comparative purposes. As seen in the figure, the achieved mean bit rates do not change significantly when comparing the SDN-based and the overflow strategy, though the SDN-based approach clearly outperforms in the less loaded situations. This is due to the fact that under high traffic loads almost all backhaul capacity (satellite and terrestrial) is being used since the UN traffic flows end up using all the available capacity and, on average, the capacity share per flow is practically the same. However, the most notorious difference comes when observing the standard deviation values, which are considerably reduced by the SDN-based strategy. This is due to the fact that this strategy distributes the traffic based on the global occupation of both satellite and terrestrial links, seeking fairness among all the established UN flows that, in the end, turns in higher network utility.

For example, under the overflow strategy we observe a substantial increase on the UN bit rates standard deviation which is about in 2.8 Mbps considering the higher UN traffic load, whereas with the SDN based strategy the same measure can be only about 1 Mbps. This is due to the fact that the strategy distributes the traffic based on the global occupation measurement of both satellite and terrestrial links, as well as the connections number present in each link.

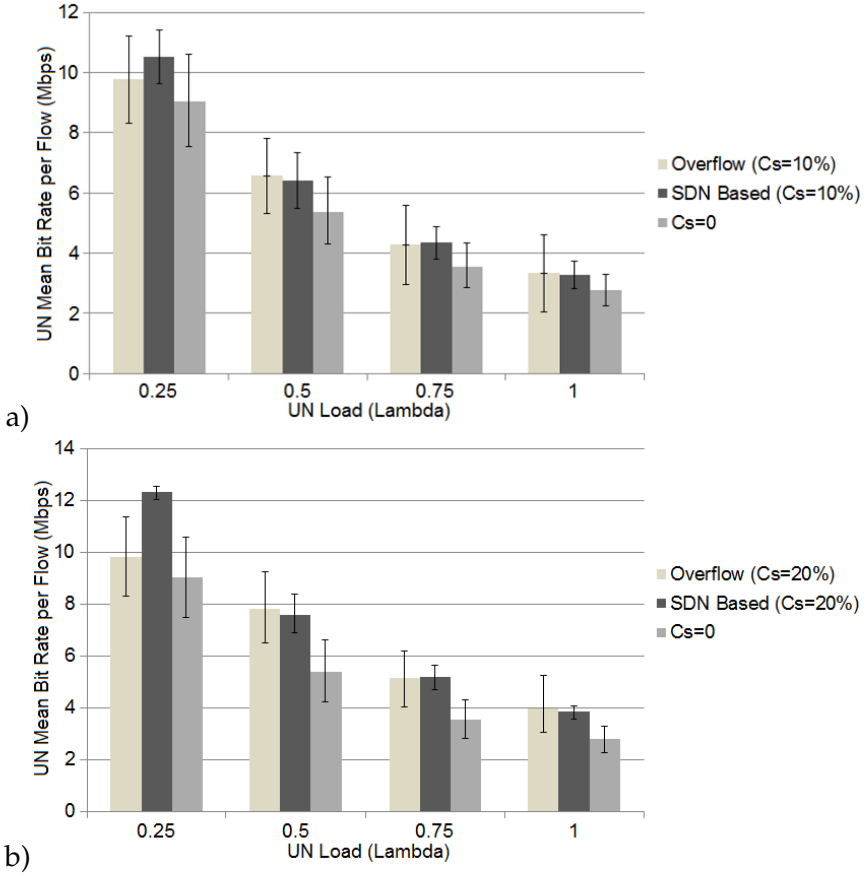

FIGURE 6.14: UN Mean Bit Rate per Flow for Cs=10% (a) and Cs=20% (b).

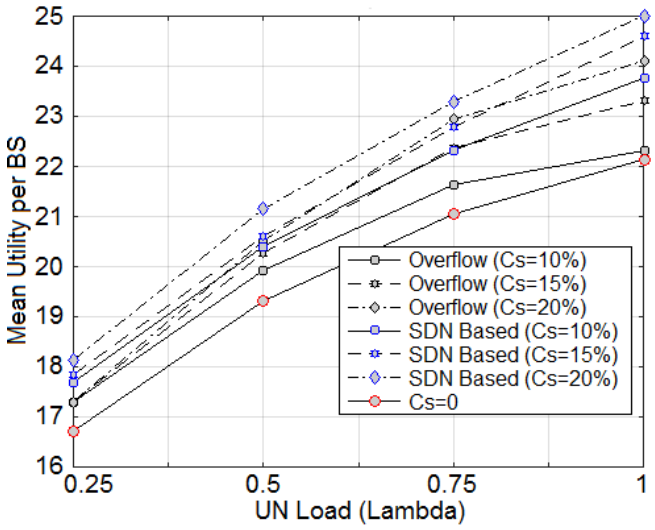

FIGURE 6.15: Utility per BS.

The network performance in terms of network utility is presented in Figure 6.15. The average utilities per BS are given in absolute terms for the SDN and Overflow strategies under different values of C_S. Results are obtained for a UG medium traffic load, $p^{UG} = p^{UN} = 1$ and the case with $C_S = 0\%$ is added for comparative purposes. It can be seen that the SDN strategy provides the highest utility in all the situations.

Deeping in details, the utility gain, computed as the percent increase of the global utility achieved by the SDN and overflow strategies with respect to that achieved for the case with $C_S = 0$ is represented on Figure 6.16. Considering $p^{UG} = p^{UN} = 1$ (a) it could be observed that, for instance, the SDN-strategy can deliver the same or even higher utility gain when operating under $C_S = 10\%$ (or 15%) than the overflow strategy for $C_S = 15\%$ (or 20%).

The higher utility achieved by the SDN strategy is partly due to the re-allocation mechanism considered as part of the TE application. In this regard, it's been assessed that the number of re-allocations that, on average, a UN flow could experience, is kept in the range of 0.26-0.65, depending on UG and UN traffic loads and showing a tendency to decrease as UN traffic increases.

Additional results in the figure show that SDN strategy is still able to bring some utility gain when considering utility reduction factors far below 1.0, e.g. gain of 4% for $C_S = 20\%$ when $p^{UG} = p^{UN} = 0.6$ (right). For factors $p^{UG} = 0.6$ and $p^{UG} = 0.6$, the same utility gain difference among strategies is slightly higher due

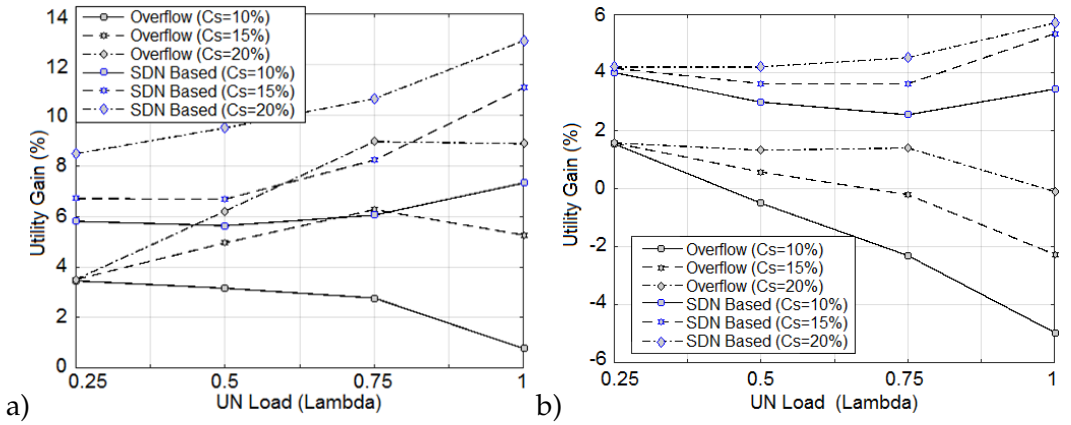

FIGURE 6.16: Global Utility Gain, with a Medium UG Traffic Load with (a) $p^{UG} = p^{UN} = 1$ and (b) $p^{UG} = p^{UN}$ =0.6.

to the overflow strategy bases the decisions only in the parameters of overflow activation/deactivation, without considering the impact that can be presented in terms of utility in some traffic type by sending connections over satellite links, being punished by a utility decrease by a smaller factor p, even achieving negative utility gains respect of not having satellite capacity.

Finally, performance results are provided considering multicast traffic. In this regard, it is assumed a MG session is forwarded, on average, to 6 BSs. Unicast traffic load is set to medium load for UG services ($\lambda_{UG} = 0.43$) and UN flows are generated with $\lambda_{UN} = 0.75$ flows/s. Multicast load is fixed as a percentage of the UG load. The satellite utility reduction factor is set to $p^{UG} = p^{UN} = p^{MG} = 0.8$ for all the services and two multicast traffic allocation strategies are considered within the SDN-based TE applications: one strategy seeks to maximize MG utility while the other strategy is intended to minimize the resource consumption of MG flows. Figure 6.17 shows the average utility achieved per BS (a) and the average mean data rate delivered to UN flows (b). As it can be observed from the figures, the strategy seeking resource consumption minimization for traffic performs much better in the two performance indicators. The reason is that resource consumption minimization enforces most MG traffic to be delivered over satellite, letting more resources available for UG and UN services that can ultimately get higher utilities and bit rates. While not reported in figures, the obtained UG average rejection rate is in the range of 0.2%-0.5% for the resource consumption minimization strategy in contrast with 0.4%-1.6% for the strategy that maximizes the MG utility.

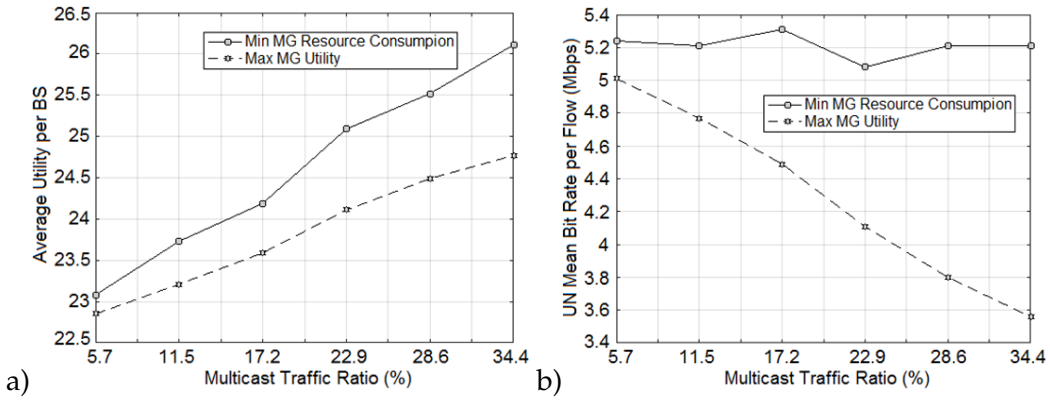

FIGURE 6.17: Multicast Traffic Handling Strategies; (a) Network Utility, (b) UN Mean Bit Rate.

6.5.4 Heterogeneous Spatial Traffic Distribution

Let us now consider the case that traffic is not homogeneously distributed among BSs. In particular, we assume that half of the BSs, denoted as group #1, are exposed to a UN load characterized with $\lambda_{UN} = (1/2) \cdot 0.75$ flows/s and, the other half, denoted as group # 2, to a UN load with $\lambda_{UN} = (3/2) \cdot 0.75$ flows/s. In addition, all the BSs support a medium UG load and the C_S is set to 20%. Under this load configuration Figure 6.18 provides the mean (a) and standard deviation (b) of the data rate achieved per UN flow. Results are given separately for the two groups of BSs, for all the BSs in the scenario and, for comparison purposes, for all the BSs under homogeneous load with $\lambda_{UN} = 0.75$ flows/s. It is observed that in group # 1 the mean bit rate provide by the overflow strategy is slightly higher than the one achieved by the SDN strategy. On the other hand, this situation is reversed for BSs in group # 2 and when considering the overall set of BSs. This mainly reflects the more fair distribution of satellite capacity enforced by the SDN-based TE application, which is even more evident when comparing the data rate standard deviation values.

Figure 6.19 presents the network performance for the different sets of BSs in terms of network utility per BS. As observed, SDN-based TE application can achieve a higher utility in the most loaded BSs (group # 2) and, as a result, a higher performance in the global scenario.

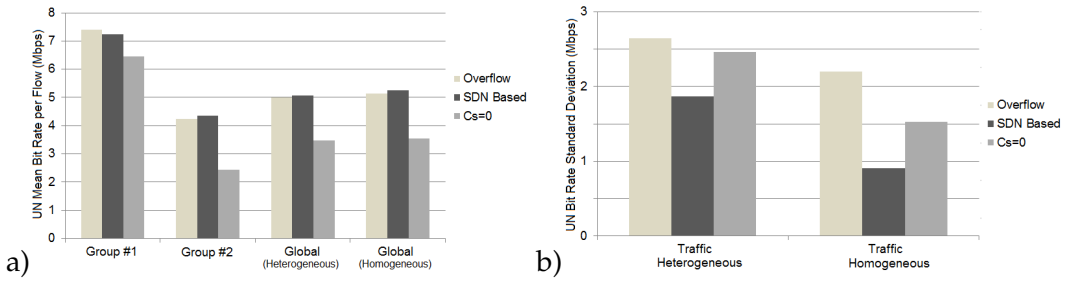

FIGURE 6.18: Mean (a) and Standard Deviation (b) of the Bit Rate Achieved per UN Flow.

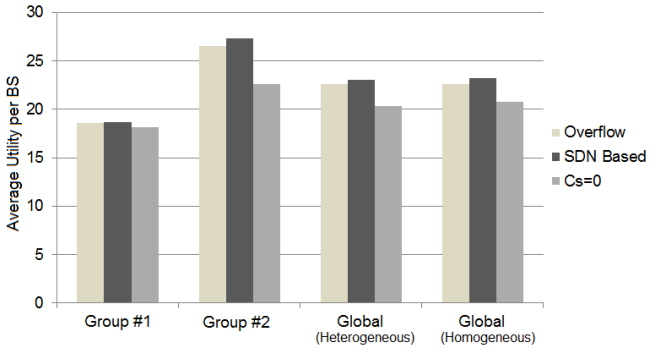

FIGURE 6.19: Average Utility per BS.

6.5.5 Satellite Backup for Terrestrial Link Failures and Transportable BSs

This second assessment is intended to show the performance of the traffic distribution strategies when not all terrestrial links are operational, especially in BSs that temporarily face a lack of terrestrial link availability, this is, when there is one BS with no terrestrial link availability in the set of $N + M = 16$ BSs that share the same pool of satellite capacity.

First, we assess the network considering one BS without terrestrial capacity, fixing the traffic load at medium UG traffic load (60%), the highest UN traffic load ($\lambda = 1$) and $C_S = 20\%$. The results are compared to those obtained when evaluating a network under the same load and where all BSs have terrestrial capacity available. Figures 6.20-6.22 show the impact in terms of utility, UG average rejection rate and UN bit rates, respectively, when one BS without terrestrial capacity is present in the network.

Results are provided separately for the BS without terrestrial capacity and for the rest of BSs in the scenario. The performance is showed for BSs where there is no terrestrial capacity (BS with no TC), in BSs where there is terrestrial capacity (BSs with

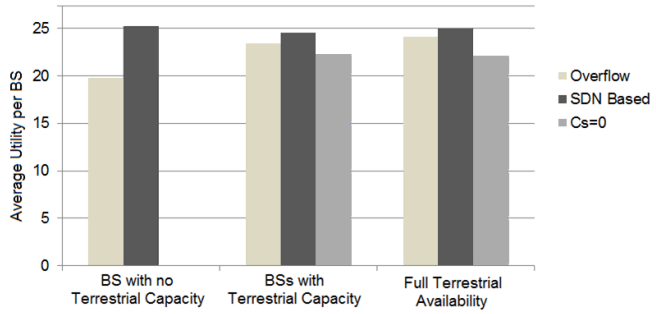

FIGURE 6.20: Average Utility for UG Medium Traffic Load, UN Flow Arrival Rate $\lambda = 1$ and $p^{UG} = 1$ and $p^{UN} = 1$.

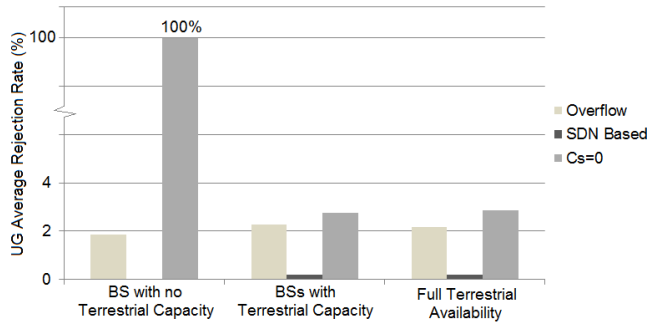

FIGURE 6.21: UG Average Rejection Rate for UG Medium Traffic Load, UN Flow Arrival Rate $\lambda = 1$ and $p^{UG} = 1$ and $p^{UN} = 1$.

TC) and results are compared with the performance reached in a network where there is terrestrial link availability in all BSs (full terrestrial availability). Moreover, for comparison purposes, each measure is compared to the case with $C_S = 0$ and the case where there all terrestrial links are fully operational.

Although the results obtained in BSs with terrestrial capacity do not show a significant impact respect to the case of a network with a full terrestrial availability, it is not so for the BS that temporarily presents a terrestrial failure. We see in Figure 6.16 that the utility values in the BS with terrestrial failure are reduced by about 20% under the conventional overflow strategy, while the SDN based strategy is able to keep the same utility levels reached in a network that does not present any terrestrial link failure. The only drawback is the number of mean UN reallocations per connection that is 0.59.

The results in terms of UG rejection rate (Figure 6.21) also show how thanks to the reservation management scheme embedded in the SDN-based TE application, the rejection rate can be fully mitigated in the BS without terrestrial capacity, while

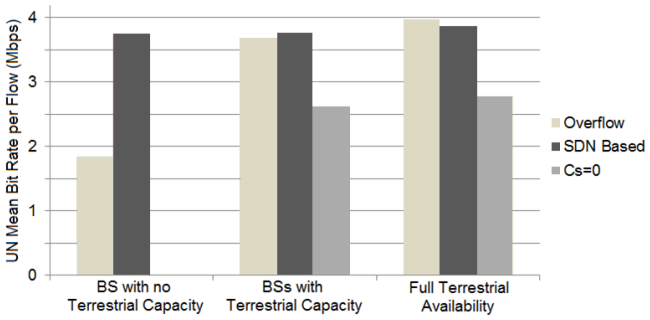

FIGURE 6.22: UN Mean Bit Rate for UG Medium Traffic Load, UN Flow Arrival Rate $\lambda = 1$ and $p^{UG} = 1$ and $p^{UN} = 1$.

under the overflow strategy the rejection rate only decreases slightly.

One of the most noticeable differences is observed in the UN mean bit rate showed in Figure 6.22. The figure presents the mean of data rates delivered to the UN flows served through the BS with no terrestrial capacity, through the rest of BSs with terrestrial capacity and in the scenario without any link failure. As seen in Figure 6.22, the UN mean bit rate achieved by the SDN-based TE application doubles that obtained under an overflow strategy in the impaired BS. This improvement is due to the applicability of the reservation management scheme in the SDN-based TE application that assures that the BS without terrestrial capacity can secure enough satellite capacity to offer an average UN bit rate comparable to the delivered through the rest of BSs. This value becomes more relevant if we take into account that under the SDN based strategy there are also no UG rejections.

Focusing now on the performance of the BS without terrestrial capacity, we assess the network under 4 different UN traffic loads ($\lambda_{UN} = 0.25, 0.5, 0.75$ and 1.0 flows/s), 3 different satellite system capacities ($C_S = 10\%$, 15%, 20%), and we fix the UG load at 60% and satellite utility reduction factor is set to 1.0 for all services. Figure 6.23 shows the utility performance of the strategies in the BS with no terrestrial capacity. The figure shows directly the utility gain in percentage obtained by the SDN strategy respect to the achieved by the overflow strategy, reaching a utility gain of up to 50% with a medium UG traffic load or even in simulations, however, for a high UG traffic load, this utility can reach up to 85%.

The UG average rejection rate is showed in Figure 6.24, showing a significant rejection rate decrease by the SDN based strategy.

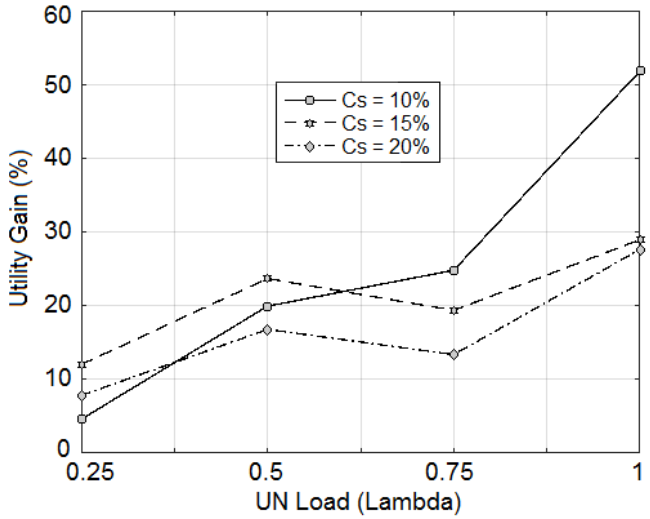

FIGURE 6.23: SDN-Based Strategy Utility Gain over Overflow Strategy at the BS with no Terrestrial Link Availability for UG Medium Traffic Load and $p^{UG} = 1$ and $p^{UN} = 1$.

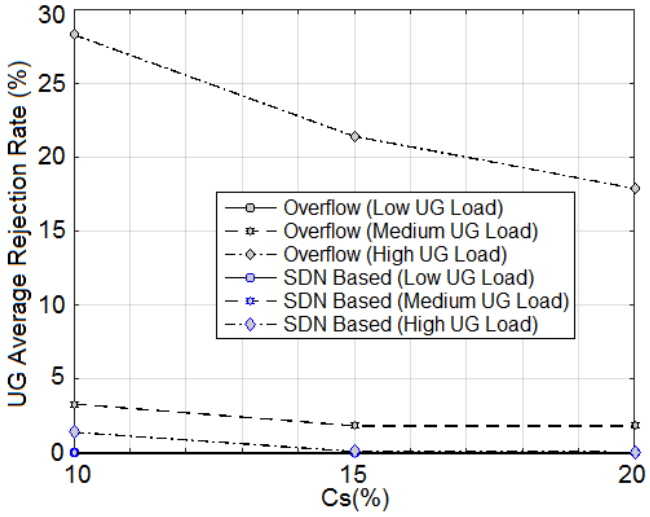

FIGURE 6.24: UG Average Rejection Rate at the BS with no Terrestrial Link Availability ($p^{UG} = 1$ and $p^{UN} = 1$).

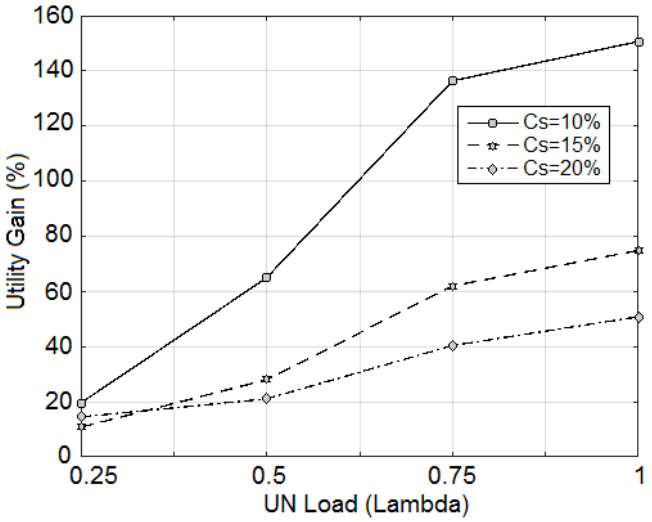

FIGURE 6.25: UN Average Utility Increase Under Failure Conditions at BS with no Terrestrial Capacity ($p^{UG} = 1$ and $p^{UN} = 1$).

We observe that in spite of accepting a larger number of UN connections, the SDN based strategy still achieves a UN utility gain of up to 150% for a medium UG load respect the achieved by the overflow strategy (Figure 6.25). The utility gains can be even greater up to 350% for a high UG traffic load. SDN based strategy has the ability to distribute all connections in such a way that a similar rate is reached among them, allowing through the reserved resources for BSs that temporarily face a lack of terrestrial capacity, a resource allocation fairness among all UN, regardless of their BS terrestrial link availability.

Likewise, Figure 6.26 shows that the SDN-based strategy can keep the UN data rates well above the overflow strategy when the UN traffic load increases.

Finally, let us consider the case that the same amount of satellite capacity provisioned to be shared in normal conditions among 16 BSs with terrestrial backhaul is used to serve a higher number of BS with no terrestrial link availability. This could be a situation where a number of transportable BSs with only satellite backhaul are brought into an incident area or the case that a disaster has severely impaired the terrestrial backhaul infrastructure of multiple BSs. Figure 6.27 provides results for up to 4 BSs with no terrestrial capacity in a group of 16 BSs. Medium UG load, UN load with $\lambda_{UN} = 1$ flows/s and $C_S = 20\%$ are considered. As expected, global network utility decreases with a higher number of BSs with no terrestrial link availability.

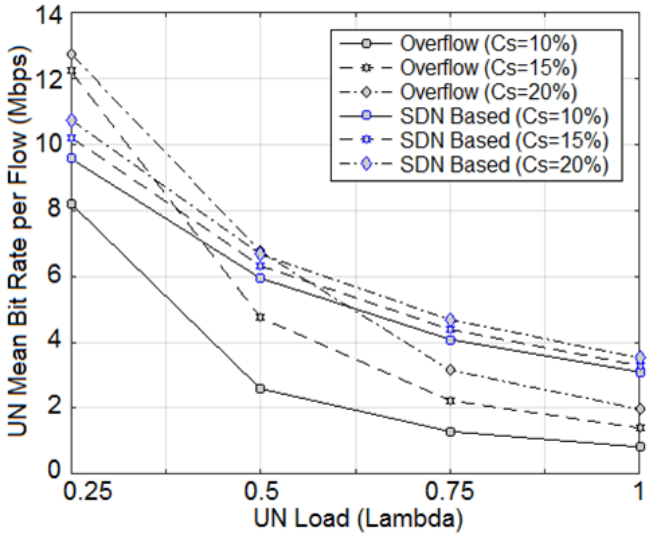

FIGURE 6.26: UN Mean Bit Rate per Flow at the BS with no Terrestrial Link Availability.

However, it could be seen that the SDN-based strategy is able to deliver higher utilities than the overflow strategy. Similar results are observed in terms of UN mean bit rates, achieving a range of 5.1-3.2 Mbps for 0 to 4 BSs with no terrestrial capacity under the Overflow strategy and of 5.3-4.2 Mbps under the SDN-based solution.

6.6 Summary

Thanks to the centralised control, the TE application allows managing the use of the satellite capacity provisioned for resilience purposes among a number of BSs so that the overall network utility is maximized under both failure and non-failure conditions in the terrestrial links. The performance of the proposed TE application has been assessed by means of numerical simulation. Obtained results shows how overall network performance is improved, compared to that of a traditional overflow solution, in terms of network utility increase, UG rejection rate decrease, UN utility increase, and improved UN fairness, especially for BSs that temporary face a lack of terrestrial link availability. Under failure conditions, it has been shown that the reservation scheme implemented within the TE application allows keeping fair utility levels in the BSs affected by terrestrial link failures.

A SDN-based TE application has been formulated that building on a global view

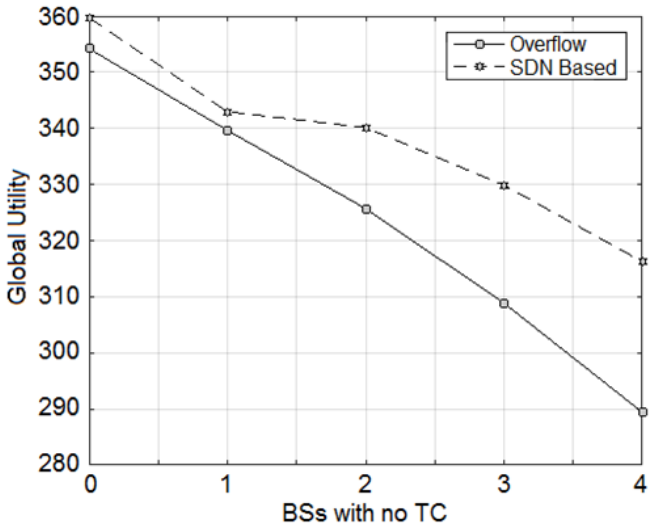

FIGURE 6.27: Global Utility for Different Number of BSs with no Terrestrial Link Availability ($p^{UG} = p^{UN} = 0.8$).

of the hybrid terrestrial-satellite network resources exploits a combination of control features and criteria such as (1) E2E path computation with terrestrial or satellite link selection; (2) satellite capacity resource reservations to deal with BSs with no or limited terrestrial link backhaul capacity; (3) different allocation criteria depending on the traffic nature (GBR and non-GBR services, unicast/multicast); (4) admission and rate control to face overload and guarantee resources and minimum (committed) transmission rates per flow and group of flows and (5) utility maximization criteria, where the adequacy of handling specific flows over the terrestrial or satellite component, as well the effect of allocating more or less data rate, are accounted. A detailed performance analysis has been conducted to assess the behavior of the proposed SDN-based TE application in multiple and diverse scenarios, including homogeneous and non-homogeneous load situations with BSs that exploit both satellite and terrestrial backhaul capacity, terrestrial link failures in some of the BSs and deployment of a number of mobile TBSs that exclusively rely on the satellite capacity for backhauling. A more traditional overflow strategy has been considered for comparison purposes. As general trends, it's been demonstrated how the proposed SDN-based TE application is able to provide a higher network utility in most of the analyzed cases, greatly improving the admission rejection ratio for GBR services and achieving higher fairness in the distribution of delivered data rates for non-GBR flows.

Chapter 7

Experimental Proof of Concept

7.1 Introduction

The functional architecture for the applicability of SDN technologies into satellite ground segment systems and its integration as part of a hybrid satellite-terrestrial network for mobile backhauling have been explained in the previous chapters. Moreover, consistently with such SDN-based functional architecture, different Traffic Engineering (TE) strategies for optimal traffic routing under diverse mobile backhaul network scenarios were formulated as SDN-based control applications and their performance assessed via numerical simulations, showing the benefits in terms of network resource efficiencies that can be achieved through the centralized, fine-grained control of traffic routing as enabled by the SDN-based architecture.

Sustained on the aforementioned SDN-based functional architecture and TE strategies and further progressing in this research area, this chapter presents an experimental Proof of Concept (PoC) and validation of the SDN-based integration solution for the satellite component in a hybrid satellite-terrestrial mobile backhaul network and the practical realization of SDN-based TE applications. In particular, building upon the functional architecture previously defined, a laboratory testbed has been developed to showcase the programming and operation of SDN-based TE applications able to enforce different traffic routing and path failure restoration policies in a hybrid backhaul network. Reported results of the PoC shed light on valuable operational data of the proposed integration solution such as the execution times of the TE mechanisms for activation and deactivation of the routing paths and the incurred signaling overhead. In addition, the PoC provides an assessment of the performance

impact caused by the applied TE policies on the QoS of the mobile network connectivity services.

The rest of the chapter is organized as follows. Section 7.2 briefly outlines the main architectural traits of the satellite-terrestrial SDN-based integration solution that is used as the conceptual foundation of this work. Next, Section 7.3 describes the implemented SDN-enabled hybrid satellite-terrestrial mobile backhaul testbed, detailing its components and service configuration settings as well as the structure and logic of the programmed E2E TE applications. On this basis, the network operational validation and performance assessment is provided in Section 7.4. Finally, Section 7.5 draws a summary.

7.2 SDN-Based Integration Solution for Hybrid Satellite-Terrestrial Mobile Backhaul

An illustrative view of the architecture of a mobile network with SDN-enabled transport from the Radio Access Network (RAN) nodes (e.g. BSs) all the way through the backhaul to the core network is depicted in Figure 7.1.

While conceptually valid for the 5G system architecture, the architectural view illustrated in Figure 7.1 is contextualized for 4G/LTE technology to facilitate the mapping with the developed testbed, which is built upon currently available LTE equipment for lab testing. As depicted in Figure 7.1, the control functions of the Mobile Core Network (MCN) (e.g. EPC control functions) are implemented as network applications that, through northbound Application Program Interfaces (APIs) provided by a SDN controller, command the operation of the data plane of the transport network (i.e. the Network Elements [NEs] with packet switching and forwarding capabilities). In this way, such SDN controller becomes instrumental for managing the connectivity services offered by the transport network (e.g. activation / deactivation of traffic flows with specific QoS settings between specific network end-points) according to the needs of the MCN applications. Moreover, the SDN controller can also be leveraged by TE applications intended to optimize the E2E transport network performance by dynamically analyzing, predicting, and regulating the traffic behavior across the network. Of note is that the SDN controller is represented here as a

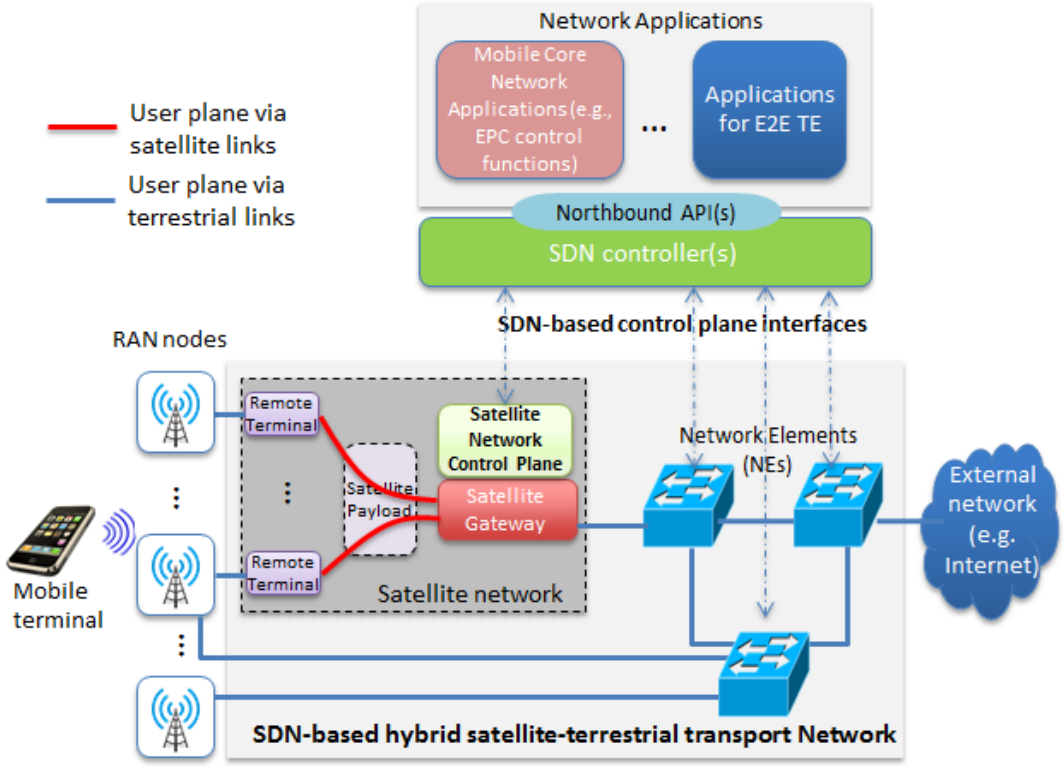

FIGURE 7.1: Illustrative View of a SDN-based Mobile Network with a Satellite Component for Mobile Backhauling within the Transport Network.

single functional entity even though, in a practical realization, it is likely to follow a hierarchical structure of controllers.

Consistently with such an SDN-based mobile network architecture defined in chapter 3 for the integration of a satellite network within the transport network that relies on the exposure by the satellite network of an SDN-based control interface for the SDN controller to manage the satellite network connectivity services as it is done with the rest of NEs. This approach is illustrated in Figure 7.1 by showing the interaction of the SDN controller with the satellite network control plane functions of the satellite network (e.g. Network Control Centre [NCC] functions in [56]). The realization of such SDN-based interface for the satellite network could actually leverage one or a combination of different interfaces and data models gaining traction in (terrestrial) networking such as OpenFlow (OF) [42], Open Networking Foundation Transport API (T-API) [47], ONF Microwave Information Model [46] and IETF Netconf along with YANG models specified for traffic engineered networks [53]. Among them, the OF specification is proposed in [115] and explained in chapter 3, to abstract the overall behavior of the satellite network as a switch, enabling in this way a fine-grained control of the forwarding behavior of the satellite network connectivity services through conventional flow management procedures [116]. Of note is that the exposure of an OF interface by a satellite network has been also proposed in the context of the realization of Virtual Network Operator (VNO) solutions [64], in which a VNO is provided with an interface to control and manage the satellite segment resources leased from a satellite network operator as if it was programming a switch OF.

Regarding the realization of TE mechanisms, the key advantage of the proposed integration solution compared to the traditional TE mechanisms used in today's transport networks (e.g. Multi-Protocol Label Switching TE [MPLS-TE]) stems from the SDN control centralization that provides a holistic view of the network together with the necessary mechanisms to enforce network polices from a single touch point. More specifically, in the hybrid satellite-terrestrial transport network scenario, such centralized control allows for the use of the satellite capacity in the way that best complements the terrestrial capacity in front of the changing conditions of both traffic demand (e.g. increase of traffic demand for an especial event, spatial demand

fluctuations over time) and network situation (e.g. backhaul backup for terrestrial link failures, network rapid roll-out, fast response capacity, cells on wheels). Indeed, the previous performance analysis shows how a given amount of satellite capacity provisioned for resilience purposes can be exploited to maximize a global network utility under both normal operation and a number of failure conditions in the terrestrial links. Such work is leveraged here for the implementation of the decision-making logic of the TE application showcased in the testbed.

7.3 Experimental Testbed

A high level view of the experimental testbed is depicted in Figure 7.2. The testbed comprises a private, small-scale LTE network that uses a SDN-enabled hybrid satellite-terrestrial network for backhauling the traffic between a BS node (called eNB in LTE) and a MCN node with the full LTE EPC functionality. The LTE network is implemented with the AMARI OTS 100 system commercialized by Amarisoft [117], which is executed over general-purpose Linux Personal Computers (PC) and Radio Frequency (RF) front-ends. In our implementation, there is one PC hosting the eNB functions and another PC executing the EPC functions (i.e. Mobility Management Entity [MME], Serving / Packet Data Network Gateway [S/P-GW] and Home Subscriber Server [HSS]). The LTE network is configured with a private network identity (i.e. a private Public Land Mobile Network [PLMN] identifier) and can be accessed by commercial LTE User Equipment (UE) such as smartphones and laptops fitted with USB data cards. Through the LTE network, UEs get IP connectivity to our laboratory network, where different application servers reside and from where access to the public Internet is also delivered.

With regard to the hybrid backhaul network between the eNB and EPC, it is built using switches OF, which stand for a particular realization of the generic NEs previously introduced in Figure 7.3. There is one switch OF at each end with two possible connectivity paths between them. One of the two paths, referred to as the terrestrial path, is implemented with a direct wired Ethernet link between the two switches OF. The other path, referred to as the satellite path, goes through a satellite link emulator implemented with the OpenSAND system [118]. The way that both paths are

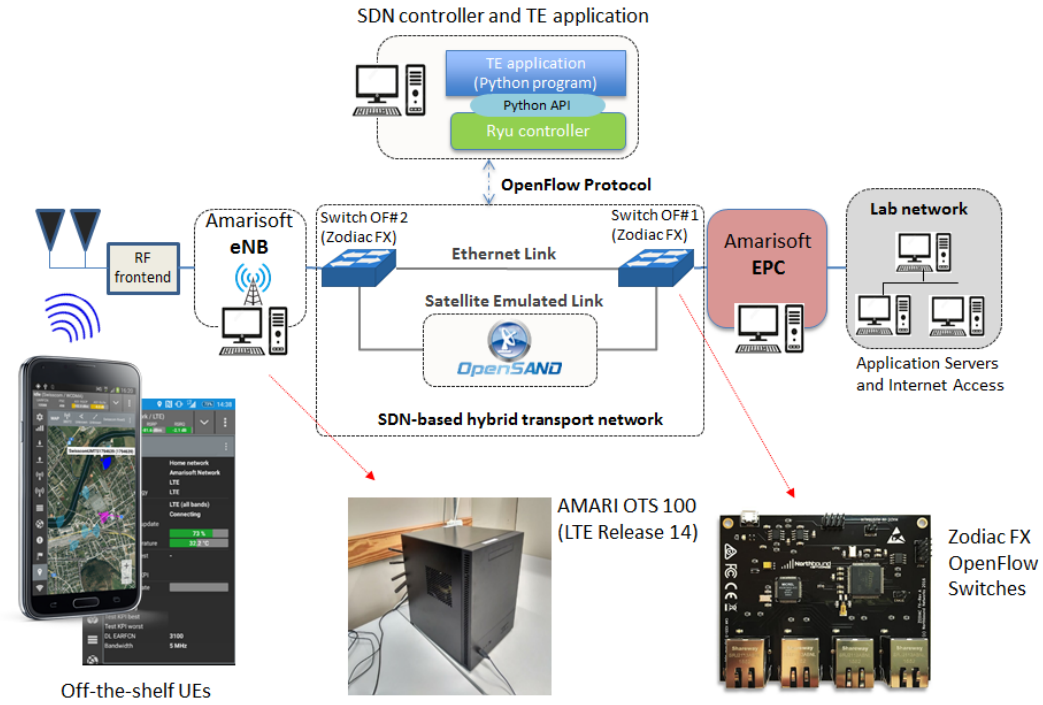

FIGURE 7.2: High Level View of the Experimental Testbed Components.

exploited is controlled by a TE application programmed in a SDN controller based on the Ryu software package [119]. A more detailed description of the testbed components is addressed in the following, including the configuration settings used to carry out the operational validation and performance assessment reported in Section 7.5.

7.3.1 LTE Network and Terminals

The private LTE network is built using the AMARI OTS 100 system [117] deployed in two separate PCs. A detailed view of the hardware elements and interconnection arrangement of the testbed components is illustrated in Figure 7.3. The two PCs running the AMARI OTS 100 system are tagged as eNB and EPC. The eNB PC is equipped with one PCIe RF board with four external antennas. The implementation is compliant with LTE Release 14, supporting the multiple frequency bands standardized below 6 GHz and cell bandwidths of up to 20 MHz. Several SIM test cards have been provisioned into the Home Subscriber Server (HSS) database of the EPC PC so that UEs can gain connectivity to the laboratory network through the private

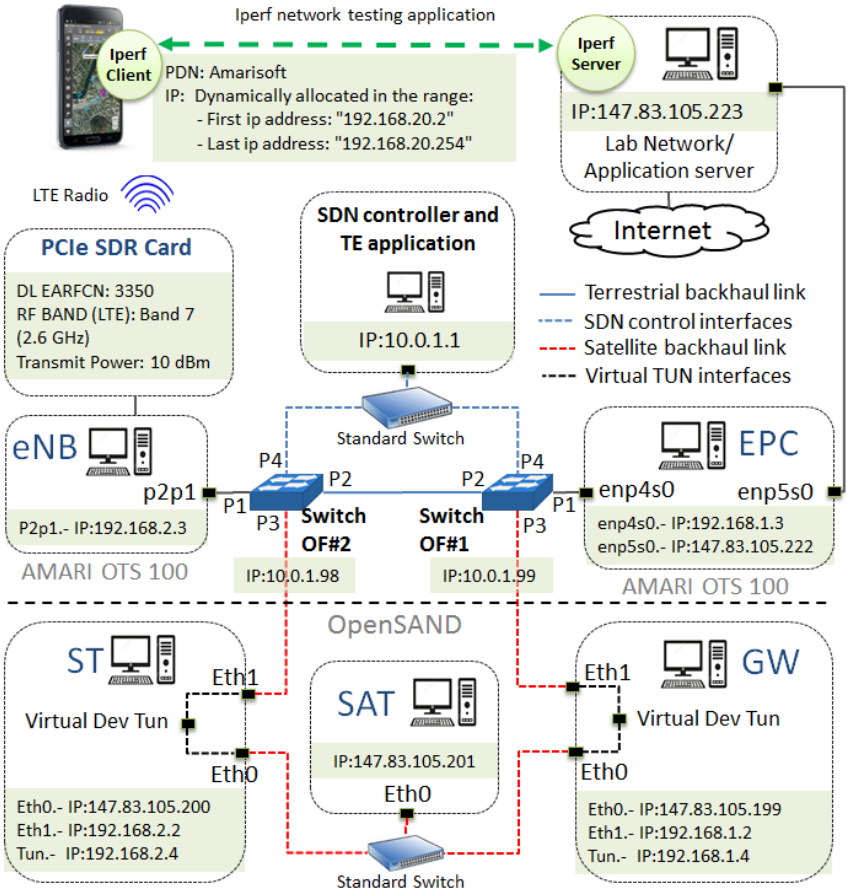

FIGURE 7.3: Detailed View of the Hardware Elements and Interconnection Settings among the Testbed Components.

LTE access. As illustrated in Figure 7.3, the eNB and EPC PCs are attached to two separate subnets (192.168.1.X and 192.168.2.X) interconnected through the switches OF. On the other hand, the pool of IP addresses assigned to the UEs belongs to another subnet (192.168.20.X) from which the lab network is reached and internet access provided through Network Address Translation (NAT) functionality in the EPC PC.

Table 7.1 provides the main configuration settings of this LTE network. Channel bandwidths of 5, 10 and 20 MHz are used in the experiments to enforce different ratios of satellite capacity over access capacity. UEs are placed at 1 meter distance from the eNB PC with line of sight to the PCI RF card. Regarding the Evolved Packet System (EPS) bearer service configuration that determines the QoS profile of IP connectivity service provided by the LTE network, the default configuration uses QoS Class Identifier (QCI) equal to 9, which corresponds to a non-guaranteed bit rate

TABLE 7.1: LTE Network Configuration

	Feature / Parameter	Description and Values
Network	Supported 3GPP standard release	Release 14
	Duplexing Mode	FDD
	Frecuency Band	LTE Band 7 (2.6 GHz)
	Frecuency Channel	DL EARFCN
	Transmit Power	<10 dBm
	Channel BW	5, 10, 20 (MHz)
	Transmission Mode (TM)	TM1
	Default EPS Bearer QCI	9 (Non-GBR)
	Default EPS Bearer ARP	15 (Lower Priority Level)
Terminals	UE	Samsung Galaxy S5
	SIM Card	symoUSIM-SJS1

(N-GBR) service typically used in the commercial networks.

7.3.2 SDN-based Hybrid Backhaul Network

The data plane of the SDN-based hybrid backhaul network includes two Zodiac FX switches commercialized by NorthBound [120]. These are experimental-oriented switches OF equipped with 4 Fast Ethernet ports, one exclusively serving as a control interface port and the rest used for traffic forwarding. In particular, as illustrated in Figure 7.3, port 4 (P4) of both switches OF is used for interaction with the SDN controller while the input/output traffic of the backhaul network is supported over P1. Using the OF protocol, the SDN controller can add, update, and delete flow entries in flow tables, both reactively (in response to new packets arriving at switches) and proactively. Each flow table in a switch OF contains a set of flow entries; each flow entry consists of match fields, counters, and a set of instructions and actions (e.g. packet forwarding actions, packet header modification actions) to apply to matching packets. Matching rules can be configured based on information such as ingress port and a combination of the multiple fields in packet headers (e.g. source/destination addresses, transport protocol, etc.). Zodiac FX switches support OpenFlow 1.3.

The emulation of a satellite link is implemented with the OpenSAND platform, which is an open-source software tool that emulates a satellite communication system based on the DVB-S2/RCS standards [118]. Our implementation of OpenSAND uses three PCs, each hosting one of three different OpenSAND modules: a satellite emulator (SAT), a gateway (GW) and a satellite terminal (ST). As depicted in

Figure 7.3, the GW and the ST PCs are directly connected, respectively, to switches OF# 1 and OF# 2 for backhaul traffic forwarding through the emulated satellite link. Within the GW and ST, the backhaul traffic is captured/injected from/to the Ethernet network cards through TUN virtual network kernel interfaces so that the be re-directed through the SAT PC through a dedicated Ethernet switch for satellite emulation processing.

Table 7.2 provides the configuration settings of the hybrid backhaul network, detailing the considered OpenSAND configuration for the emulation of the forward and return link transfer rates and one-way latencies for different tested satellite types (e.g. GEO, MEO). The capacity of the satellite link has been defined according to the values of typical capacities offered commercially [121]. Likewise, the values of the latencies have been obtained from the parameters offered by current commercial systems, GEO and MEO (O3b system) shown in [122].

TABLE 7.2: SDN-based Hybrid Backhaul Network Configuration

	Feature / Parameter	**Values**
Zodiac FX Switches	Operation Mode	OpenFlow 1.3
OpenSAND	System type	DVB-S2/RCS
	Forward Link Symbol Rate	20 E6, 28.8 E6 bauds
	Forward Link Modulation	QPSK
	Forward Link Coding Rate	1/4, 2/3
	Forward Link Capacity	10, 30 Mbps
	Return Link CRA[a]	10 Mbps
	Return Link Max RBDC[b]	1024 Kbps
	Return Link Max VBDC[c]	55 Kbps
	Return Link Capacity	10 Mbps
	One-way latency	250 ms (for GEO at 36,000 Kms) 70 ms (for MEO at 8,000 Kms)

[a]Constant Rate Allocation.
[b]Rate Based Dynamic Capacity. Assigned dynamically at the request of the Return Channel Satellite Terminal (RCST) according to the standard DVB-RCS.
[c]Volume Based Dynamic Capacity. Assigned dynamically at the request of the Return Channel Satellite Terminal (RCST) according to the standard DVB-RCS.

7.3.3 SDN-based TE Application

The operation of the switches OF is commanded by a Ryu SDN controller, which is a component-based software fully written in Python. The Ryu controller exposes Application Programming Interfaces (API) for deploying network management and

control applications as Python scripts [116]. Such APIs are a collection programming libraries that give access to the previously mentioned set of mechanisms supported by OpenFlow 1.3 so that capabilities for network state monitoring (e.g. switch status, port status, traffic load) and activation/ deactivation/ modification of flow tables in the switches OF are visible at application level.

The exposed API capabilities, allows us to program and demonstrate the operation of an SDN-based TE application able to deliver the sort of TE strategies studied in chapter 5-6. More specifically, the implemented SDN-based TE application is able to (1) learn the network topology, (2) monitor the network/port status, (3) detect new traffic at input/output ports of the backhaul network, (4) decide on the forwarding path based on a set of user-defined path computation policies and (5) enforce the desired forwarding path by populating the Flow Tables across the switches OF. An illustration of the implemented SDN-based TE application is given in Figure 7.4, showing its internal organization and the exploited APIs. Of note is that the path computation policies are defined as a set of vectors where each row defines (1) Traffic matching conditions, including the traffic flow templates (e.g. discrimination of TCP, UDP, SCTP and/or ports of these protocols) and the input port of switches OF where the traffic flow is detected; (2) the network status of the switches OF ports (e.g. Up/Down); and (3) the preferred path policy (e.g. terrestrial or satellite path).

7.4 Operational Validation and Performance Assessment

The operation of the testbed is validated through the execution of two illustrative TE procedures. The first one, referred to as path computation, shows how the implemented SDN-based TE application is able to enforce a desired routing over the hybrid backhaul each time a new LTE connection is to be established. On the other hand, the second procedure, called hereafter path restoration, shows how an unexpected failure of a link is handled by the SDN-based TE application so that the affected traffic flows are quickly re-directed over an alternative path. A description of the signaling flows and key decision-making points in the execution of the two procedures is firstly covered in this section. Next, two performance assessments are provided to complement the operational validation: one showing the impact on the

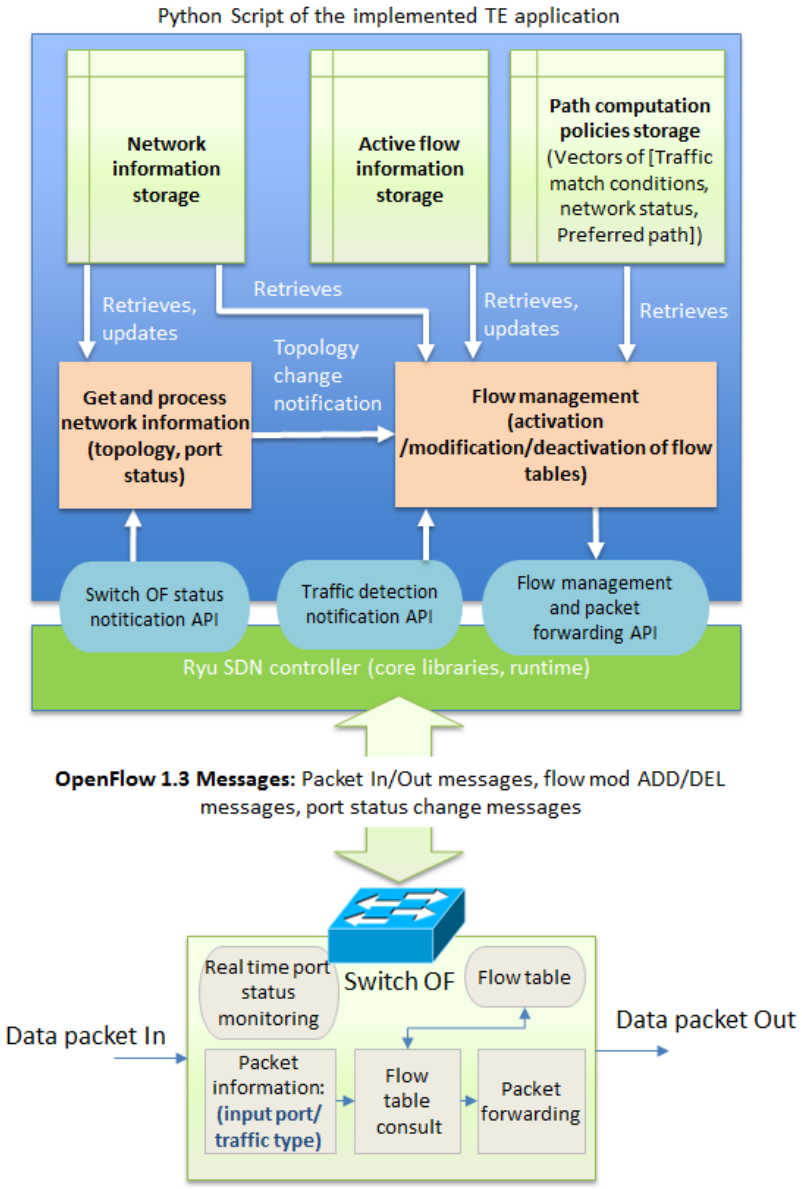

FIGURE 7.4: Illustration of the Implemented SDN-based TE Application.

LTE service performance of the routing decisions by the TE application and another addressing the impact of the link failure and path restoration procedure.

7.4.1 Execution of Path Computation and Restoration

For the demonstration of the path computation procedure, the testbed is started with the configuration explained in the previous section and a LTE UE is switched on and registered in the LTE network. This results in the activation of the so called default EPS bearer service in the LTE network, which provides the basic IP connectivity service to the UE. Then, right after getting the connectivity service, an Iperf client application [123] is launched in the UE so that UDP traffic begins to flow between the UE and an Iperf server located in the laboratory network. The routing of Iperf generated traffic through the hybrid backhaul is enforced by the TE application in the SDN controller based on the detection of the traffic type supported by the bearer service, the port where this traffic is detected and the network status. In particular, the path computation policy established for this scenario consists in forcing all traffic between the eNB and EPC (i.e. match condition: Traffic Type=*Any*, Input Port=P1 at both OF# 1 and OF# 2) to go through the terrestrial link (i.e. output port=P2 at both OF# 1 and OF# 2) whenever the backhaul network is operating under normal conditions (i.e. P1/P2/P3 status=Up at both OF# 1 and OF# 2 switches). Otherwise, when one of the paths is not available (i.e. P2 or P3=*Down* at OF# 1 or OF# 2), all the established flows are routed through the alternative path.

The results of the execution of the test are illustrated in Figure 7.5 by showcasing the signaling exchanges between the testbed components with particular focus on the steps from the detection of the Iperf application traffic by the TE application at the SDN controller to the point at which the necessary flows are established in the switches OF for proper traffic forwarding through the backhaul links. All this information is collected with the help of the management tools of the AMARI OTS 100 and Zodiac FX switches and by inspecting traffic on the SDN controller, eNB and EPC with a network protocol analyzer tool (Wireshark).

The first step shown in Figure 7.5 corresponds to the initial OF signaling handshaking between the SDN controller and the switches OF for the activation of the control plane. This signaling goes through the P4 ports and consequently does not

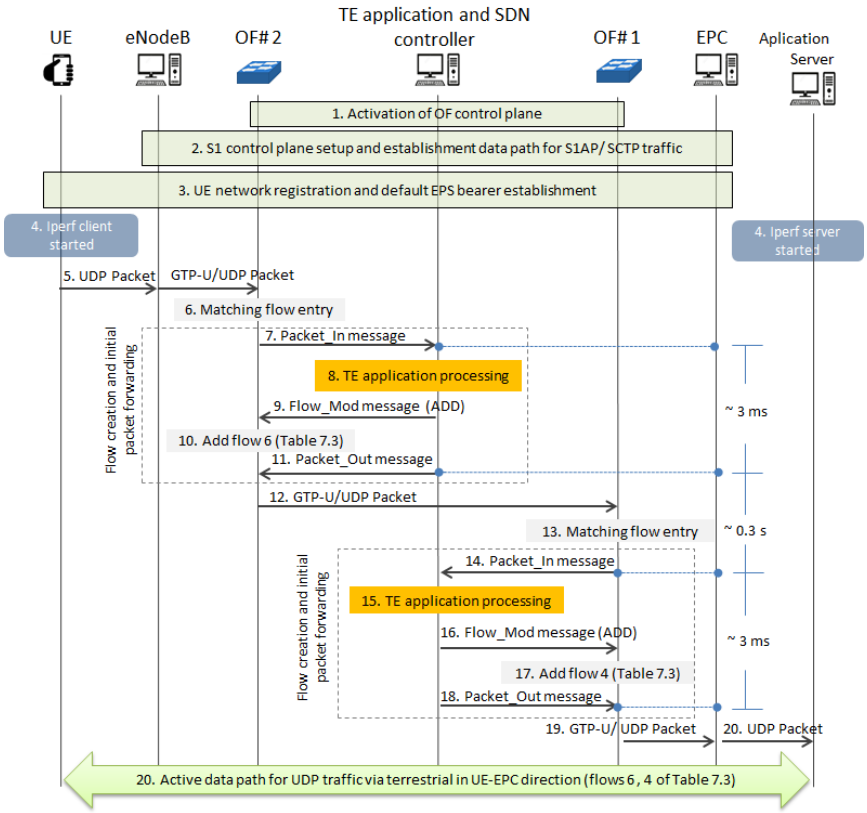

FIGURE 7.5: Signalling Exchange for the Path Computation Procedure.

traverse the hybrid backhaul data plane. The second step corresponds to the initial signaling exchange between the eNB and the EPC. This signaling, commonly denoted as S1 Setup and which uses S1 Application Protocol (S1AP) and Stream Control Transmission Protocol (SCTP) as transport protocol, is necessary for the activation of the control plane of the S1 interface between the eNB and the EPC components of the LTE network. Unlike the previous mentioned OF signaling, now the S1AP signaling is indeed transferred over the hybrid backhaul network, so that the TE application detects this traffic in both directions and creates the corresponding flows (i.e. flows 1,3,5 and 7 in Table 7.3). Next, once the S1 interface is operational, the step 3 in Figure 7.5 corresponds to the network registration and activation of the default EPS bearer service when the UE is switched on. In this case, the signaling originated in step 3 between the eNB and EPC is sent through the path previously established in the activation of the control plane of S1, so that no new flows are created at this point. Reached this situation, the Iperf traffic is started at step 4. This result in a new IP packet detected at OF# 2 for which no matching flow entry exists (step 5 and 6). The detected packet is indeed a GPRS Tunneling Protocol for User Plane (GTP-U) packet that encapsulates the original UDP packet received from the UE. As a result, the TE application is notified in step 7 of the presence of new traffic. This triggers the computation of the routing path at step 8, where, according to the established path routing policy, the creation of a new flow (flow 6 in Table 7.3) is decided. This decision triggers the activation of the flow through the OF protocol signaling (step 9 and 10) together with a specific OF command (step 11) given to the switch to forward the packet that triggered the flow creation. Once the forwarded GTP-U/UDP packet reaches OF# 1, the same process of creation of a flow in OF# 2 is now repeated for OF# 1 (steps 14 to 18), leading to the creation of a new flow (flow 4 in Table 7.3) and to the eventual forwarding of the GTP-U/UDP packet towards the EPC (step 19). Finally, the original UDP packet is decapsulated in the EPC and, after NAT processing, forwarded to the Iperf server within the laboratory network. At this point, the connectivity path in one direction is fully operational between the eNB and EPC so that subsequent packets will already match the created flow entries and therefore be forwarded without requiring any interaction with the SDN controller.

Later, when traffic is detected in the opposite direction, the process is repeated concluding with the creation of the flows 2 and 8 in 7.3. Figure 7.5 also shows that the overall processing time for flow creation Table 7.3: Flows Created During the Path Computation and Restoration Procedures including the decision-making process of the TE application is roughly ∼3 ms measured from the arrival of the new traffic notification (step 7) to the sending of the output packet command (step 11). On the other hand, the response time between the eventual traffic forwarding at OF# 2 until the notification of the detection of the packet at OF# 1 (step 14) results in the order of ∼300 ms. Of note is that this response time is determined by the switch manufacturer design, which is to be much lower in a carrier-grade product. The flows created for traffic S1-AP/SCTP and GTP-U during the whole Path Computation process are summarized in Table 7.3, totaling 8 different flows (4 at each switch OF).

TABLE 7.3: Flows Created During the Path Computation and Restoration Procedures

Procedure	Switch	Flow	In Port/Out Port	Traffic Type
Path Computation	OF#1	1	P1 / P2	S1-AP/SCTP
		2	P1 / P2	GTP-U/UDP
		3	P2 / P1	S1-AP/SCTP
		4	P2 / P1	GTP-U/UDP
	OF#2	5	P1 / P2	S1-AP/SCTP
		6	P1 / P2	GTP-U/UDP
		7	P2 / P1	S1-AP/SCTP
		8	P2 / P1	GTP-U/UDP
Path Restoration	OF#1	9	P1 / P3	S1-AP/SCTP
		10	P1 / P3	GTP-U/UDP
		11	P3 / P1	S1-AP/SCTP
		12	P3 / P1	GTP-U/UDP
	OF#2	13	P1 / P3	S1-AP/SCTP
		14	P1 / P3	GTP-U/UDP
		15	P3 / P1	S1-AP/SCTP
		16	P3 / P1	GTP-U/UDP

For illustrative purposes, Figure 7.6 shows a Wireshark capture at the PC running the SDN controller with the signalling messages originated between Steps 7, 9, 11, 14, 16 and 18 of message chart of Figure 7.5. With the same purpose, Figure 7.7 shows the information provided by the Zodiac FX management tool regarding to the configuration of one of these flows.

For the demonstration of the path restoration procedure, the testbed is started

Time	Source	Destination	Protocol	Length	Info
17.6739292	10.0.1.98	10.0.1.1	OpenFlow	214	Type: OFPT_PACKET_IN
17.6765547	10.0.1.1	10.0.1.98	OpenFlow	174	Type: OFPT_FLOW_MOD
17.6769505	10.0.1.1	10.0.1.98	OpenFlow	228	Type: OFPT_PACKET_OUT
17.9636365	10.0.1.99	10.0.1.1	OpenFlow	206	Type: OFPT_PACKET_IN
17.9663060	10.0.1.1	10.0.1.99	OpenFlow	174	Type: OFPT_FLOW_MOD
17.9666130	10.0.1.1	10.0.1.99	OpenFlow	220	Type: OFPT_PACKET_OUT

FIGURE 7.6: OF Messages Originated in the Steps 7,9,11 and 14,16,18 Illustrated in Figure 7.5.

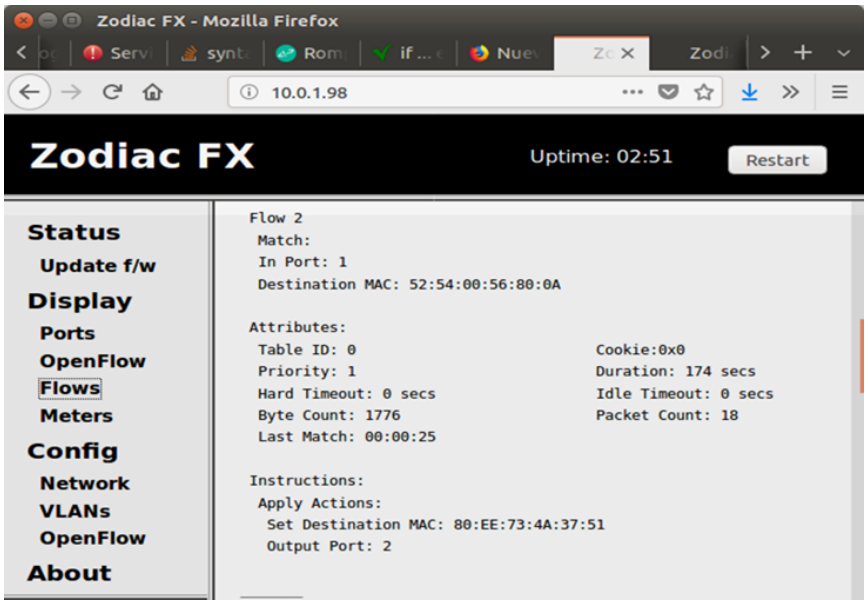

FIGURE 7.7: Illustration of the Flow Attributes Information as Displayed by the Zodiac FX Switches Management Tool.

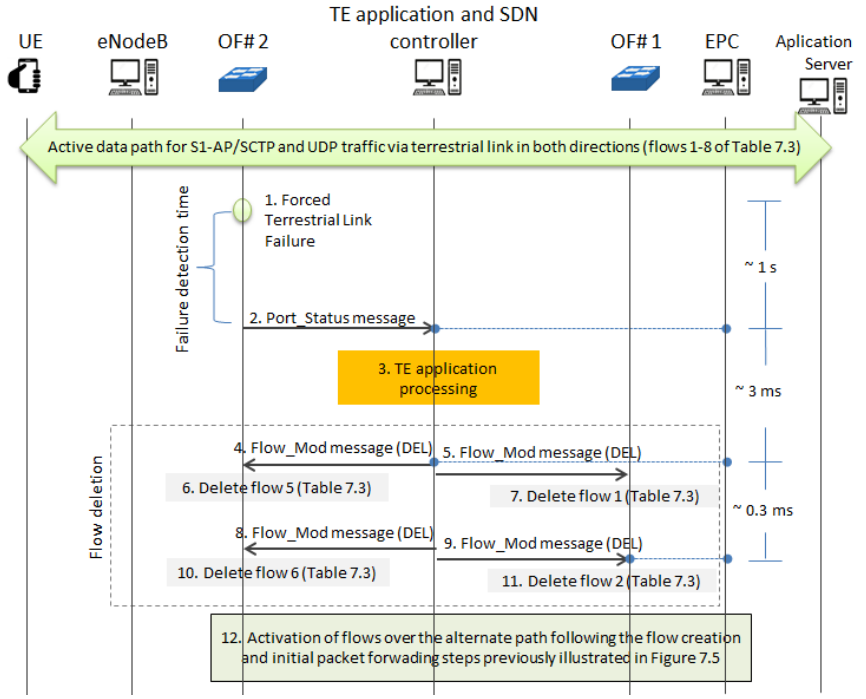

FIGURE 7.8: Signalling Exchange for the Path Restoration Procedure.

as in the previous demonstration and, once the connectivity is operational and the Iperf traffic is flowing, a link disconnection is manually forced. Figure 7.8 shows the signalling originated as a result of this disconnection when the affected path is the terrestrial one.

As seen in Figure 7.8, the link failure event is detected at OF# 2 that changes the status of P2 (from *Up* to *Down*) and sends an OF message (Step 2) to the SDN controller (it should be noted that in this case the link event is detected at both switches OF, so the port status change notification [step 2] actually arrive from both switches OF, even though Figure 7.8 only shows the link failure event detected only at OF# 2). This triggers a notification to the TE application that checks which flows at both switches OF are affected by the disconnection (i.e. all flows that have as output port = P2). Accordingly, the TE application triggers the deletion of these flows between steps 4 and 11 (i.e. flows 1, 2, 5 and 6 in Table 7.3) so that traffic forwarding towards the impaired path is stopped. In terms of execution times, Figure 7.8 shows a response time in the order of 1 second between the link disconnection and the reception of the port status message at the SDN controller, which is again attributable

Time	Source	Destination	Protocol	Length	Info
32.5654266	10.0.1.98	10.0.1.1	OpenFlow	134	Type: OFPT_PORT_STATUS
32.5682950	10.0.1.1	10.0.1.98	OpenFlow	134	Type: OFPT_FLOW_MOD
32.5684106	10.0.1.1	10.0.1.99	OpenFlow	134	Type: OFPT_FLOW_MOD
32.5686188	10.0.1.1	10.0.1.98	OpenFlow	294	Type: OFPT_FLOW_MOD
32.5686310	10.0.1.1	10.0.1.99	OpenFlow	294	Type: OFPT_FLOW_MOD
32.5688950	10.0.1.99	10.0.1.1	OpenFlow	134	Type: OFPT_PORT_STATUS

FIGURE 7.9: OF Messages Originated in the Steps 2,4,5,8 and 9 Illustrated in Fig. 8.

to the specific hardware used in the testbed. On the other hand, the process involving the TE application decision-making and the deletion of the flows remains in the order of a few milliseconds.

After the removal of the flows, the establishment of the new path is solved in the implemented TE application by reactively triggering the creation of new flows when notifications about packets not finding a flow matching entry are received in the SDN controller. This process is indeed the same as the one followed in the path computation procedure previously explained in Figure 7.5. The main difference is that now network conditions are different (i.e. P2 is down) so that the TE application decides creating the new flows using the satellite path. The new flows created in this case are listed as flows 9 to 16 in Table 7.3.

For illustrative purposes, Figure 7.9 shows a Wireshark capture of the message sequence from the reception of port status change notification message from OF# 2 at the SDN controller, followed by four OFPT_ FLOW_ MOD messages sent by the SDN controller to the switches OF# 1 and OF# 2 in order to delete the previously referred flows. It's worth noting that the OF protocols uses the same type of OFPT_ FLOW_ MOD messages for both creation and removal of flows, though internal headers define whether it is an OFPFC_ ADD or OFPFC_ DELETE operation, as illustrated in Figure 7.10.

7.4.2 Impact of the Path Selection on the LTE Service

The impact on the LTE service performance of selecting the connectivity path via the terrestrial or satellite link is evaluated by assessing at application level the experienced (1) round-trip time (RTT), (2) maximum TCP throughput and (3) maximum

a)
```
▾ OpenFlow 1.3
    Version: 1.3 (0x04)
    Type: OFPT_FLOW_MOD (14)
    Length: 120
    Transaction ID: 1662844462
    Cookie: 0x0000000000000000
    Cookie mask: 0x0000000000000000
    Table ID: 0
    Command: OFPFC_ADD (0)
```

b)
```
▾ OpenFlow 1.3
    Version: 1.3 (0x04)
    Type: OFPT_FLOW_MOD (14)
    Length: 80
    Transaction ID: 1305703908
    Cookie: 0x0000000000000000
    Cookie mask: 0x0000000000000000
    Table ID: 0
    Command: OFPFC_DELETE (3)
```

FIGURE 7.10: Flow_ Mod Message Commands; (a) to Create OFPFC_ ADD and (b) to Delete OFPFC_ DELETE.

UDP throughput achievable through the LTE network. In particular, the RTT is measured with the ping utility between the UE and the application server and the Iperf tool is used for throughput characterization. For UDP traffic, the value reported as maximum throughput corresponds to the highest throughput reached with less than ∼1 datagram loss. Average values are computed from the execution of 10 tests of 100 seconds duration each.

Table 7.4 provides obtained results under 3 different configurations of the LTE channel bandwidth (i.e. 5 MHz, 10 MHz and 20 MHz) and 4 different configurations of the satellite link (i.e. forward link capacity of 10 or 30 Mbps and, in each case, consideration of GEO or MEO latency characteristics). Focusing first on the cases where the traffic is forwarded through the terrestrial link, it is seen that RTT values are of the order of 35 ms regardless of the LTE channel bandwidth and that achieved throughputs vary from roughly 15 to 70 Mbps depending on both the traffic nature and channel bandwidth. This performance is clearly limited by the LTE access network itself since the backhaul path in this case is a direct Ethernet connection, which is far from becoming the bottleneck under this configuration. On the other hand, focusing now on the performance values achieved when the traffic goes through the satellite path, the impact of the different configurations of the satellite link are now noticeable if benchmarked against the performance values achieved in the previous case. With regard to RTT, values now have increased from 35 ms up to around 230 ms or 580 ms for MEO and GEO type links respectively, which are clearly dominated by the latency performance of the satellite component.

On the other hand, the impact of the satellite link on the maximum achievable

UDP throughput is only noticeable in the situations that the provisioned satellite capacity is below the achievable LTE access capacity. Hence, when LTE access is using a 5 MHz channel, a forward satellite link capacity of 10 Mbps originates a throughput reduction from 17.5 Mbps to 8.6 Mbps while a forward satellite link capacity of 30 Mbps does not bring any impairment. As a general observation, it could be seen that UDP throughput limitation is basically a matter of capacity limitation.

TABLE 7.4: Impact of Path Computation Decisions on LTE Service Performance.

				LTE Channel Bandwidth		
				5 MHz	**10 MHz**	**20 MHz**
TE routing: Terrestrial Path			RTT	30.63 ms	36.12 ms	35.11 ms
			TCP	16.5 Mbps	34.7 Mbps	71.9 Mbps
			UDP	17.5 Mbps	35.9 Mbps	59 Mbps
TE routing: Satellite Path	Forward Link Cap.:10 Mbps	MEO	RTT	232.5 ms	231.6 ms	239.8 ms
			TCP	7.37 Mbps	7.56 Mbps	7.69 Mbps
			UDP	8.6 Mbps	8.6 Mbps	8.6 Mbps
		GEO	RTT	591.1 ms	581.9 ms	590.1 ms
			TCP	7.06 Mbps	7.38 Mbps	7.51 Mbps
			UDP	8.6 Mbps	8.6 Mbps	8.6 Mbps
	Forward Link Cap.:30 Mbps	MEO	RTT	229.1 ms	236.7 ms	234.7 ms
			TCP	7.1 Mbps	6.46 Mbps	7.17 Mbps
			UDP	17.4 Mbps	27.2 Mbps	27.2 Mbps
		GEO	RTT	579.7 ms	577.8 ms	580.1 ms
			TCP	5.67 Mbps	5.53 Mbps	5.40 Mbps
			UDP	17.4 Mbps	27.2 Mbps	27.2 Mbps

The situation gets more complex when it comes to TCP traffic. In this case, when considering a satellite capacity of 10 Mbps, the maximum throughput rate remains around 6.5-7.6 Mbps, which is expectedly limited by the satellite capacity. However, when considering a higher satellite link capacity (30 Mbps), the throughput achieved through the LTE network does not improve under the MEO configuration, yielding around 7 Mbps irrespective of the LTE channel bandwidth, and even gets worse for the GEO case, yielding around 5.5 Mbps also regardless of channel bandwidth. A more detailed analysis of this behavior, out of the scope of this paper, has revealed that the low TCP throughputs are actually due to a high fluctuation observed in the instantaneous TCP throughput, with sharp drops that bring down the averaged values.

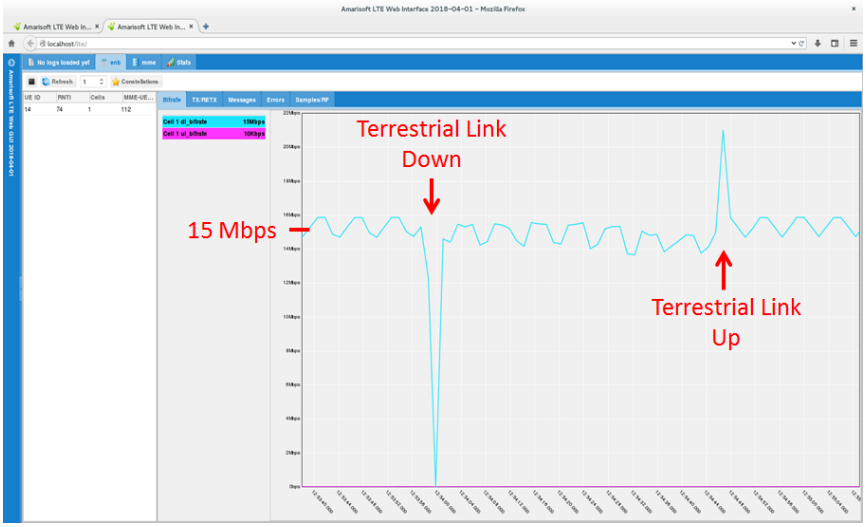

FIGURE 7.11: Time Evolution of the UDP Throughput During the Test, Highlighting the Points at which Terrestrial Link Goes down and up Again.

7.4.3 Impact of the Path Restoration Procedure on the LTE Service

Finally, the impact on the LTE service performance of a backhaul link failure and the subsequent execution of the path restoration procedure is analyzed in this section. To that end, the LTE channel bandwidth is set to 10 MHz and OpenSAND is configured with a forward link capacity of 30 MHz and GEO-type latency. When the testbed is started, the Iperf application is used to transfer UDP traffic at 15 Mbps. Such traffic is initially routed through the terrestrial path and, at a given point in time, a terrestrial link failure is provoked and the TE application takes care of re-routing the UDP traffic through the satellite path following the path restoration procedure previously analyzed. Then, after a few seconds of sending the traffic over the satellite, the terrestrial link is turned up again and the TE application triggers again a path restoration procedure now to move back the UDP traffic over the terrestrial path. An illustration of the tested behavior is shown in Figure 7.11 as measured and visualized from the Amarisoft management tool at the LTE eNB. The same experiment has been with UDP transfer rates of 5, 15 and 25 Mbps, each executed 10 times for statistics gathering.

Focusing first on the link failure event, it could be seen in Figure 7.11 that the time taken for the path restoration procedure causes a noticeable throughput drop. A quantification of this traffic drop is provided in Table 7.5 in terms of the number

of datagrams lost for 5, 15 and 25 Mbps UDP traffic rate. As it can be observed from the table, the amount of datagrams lost is directly proportional to the bit rates and the path restoration time, which is of the order of 1 second in our setting due to the response time of the utilized OF switching hardware. Therefore, the impact due to the processing time associated with the operation of the TE application is almost negligible. As a side note to further stress the advantage of the implemented TE application, turning the operation mode of OF# 1 and OF# 2 from "OpenFlow" to "standard", that is, letting the re-direction of traffic in the backhaul to be fixed by the conventional distributed spanning tree protocol supported in the switches, and repeating the same experiment, observed interruption times boost above tens of seconds.

Moving now to the point that the terrestrial path goes operational again and the traffic is switched back to this path by the TE application, no packet losses are observed in this case. This is because the traffic can be still correctly transferred through the satellite link which remains operational during the path switching process. Remarkably, this situation even results here in a temporary increase of the throughput due to the simultaneous arrival of traffic via both links. However, since the satellite link latency is higher, the first packets sent through the terrestrial link are out of order. A quantitative assessment of this effect is also presented in Table 7.5, where it can be seen that the amount of datagrams received out of order is proportional to the UDP application bit rate and the satellite latency.

TABLE 7.5: Impact of the Path Restoration Procedure on LTE Service Performance.

Terrestrial Link Transitions	Iperf Bit Rate (Mbps)/ Datagrams per sec	Datagrams Lost (Avg./σ)	Datagrams received out of order (Avg./σ)
Up to Down	5 / 425	465 / 211	-
	15 / 1275	1566 / 674	-
	25 / 2125	2150 / 615	-
Down to Up	5 / 425	-	99.7 / 5.69
	15 / 1275	-	204.2 / 55.8
	25 / 2125	-	372.2 / 4.86

7.5 Summary

The adoption of SDN technologies is nowadays acknowledged as a key enabler for a better integration of a satellite component into 5G networks. Sustained in the architectural defined in the chapter 3, this chapter has presented an experimental PoC and validation of an integration solution based on the use of SDN technologies for the realization of E2E TE applications in hybrid terrestrial-satellite backhaul mobile scenarios.

The implemented testbed has allowed us to assess the feasibility of the proposed SDN-based integration solution under a practical laboratory setting that combines the use of commercial, experimentation-oriented and emulation equipment and software. Procedures for path computation and restoration have been showcased, assessing the incurred execution times and signaling exchanges incurred in the different steps. In particular, It has been demonstrated that the time incurred in the creation and deletion of flows from the controller is of the order of a few milliseconds, and that the longest delays incurred in the tested procedures that actually have a noticeable impact on service performance are mainly associated with the nature of the hardware used in the experimentation. Moreover, it has been also validated the versatility of using a high-level language like Python and the existing OF libraries in the Ryu controller for programming the TE application in order to enforce different traffic routing and path failure restoration policies. Finally, the testbed has also allowed us to assess the impact on the LTE service performance for TCP and UDP traffic flows when using and switching to/from a satellite link.

Chapter 8

Conclusions and Contributions

8.1 Introduction

This chapter concludes this dissertation presenting in a summarized way the set of contributions made throughout this research work. It gathers the most important conclusions as part of the obtained results. Likewise, an analysis is presented on the future work that could contribute complementary in line to this research work.

The motivation behind this work lies in that despite the adoption of SDN technologies is acknowledged as a key enabler for a better integration of a satellite component into 5G networks, and the existence of a growing consensus among the industry and standardization bodies as the 3rd Generation Partnership Project (3GPP), on the idea that satellite could complement terrestrial 5G network components in a number of scenarios, little attention has been paid to capacity and traffic management aspects (e.g., capacity dimensioning, routing, QoS, congestion, resilience) when it comes to the consideration of hybrid terrestrial-satellite networks for mobile backhauling. The next section summarizes the main contributions of the thesis.

8.2 Contribution

This thesis has delineated the potential applicability of a satellite networks as part of a hybrid satellite-terrestrial mobile backhauling network and described a view for the seamless integration of the satellite component that is aligned with the current evolution of mobile and transport networks towards SDN architectures.

The thesis first has studied and described an architectural framework that enables the integration and management of the satellite capacity as a constituent part

of a Software Defined Networking (SDN) -based traffic engineered mobile backhaul network for a better provisioning and operation of the satellite component in a more flexible, agile and cost-effective manner than done today.

With regard to architectural aspects, building upon the ETSI functional architecture for BSM systems, this thesis has made an analysis of a solution for the adoption of an SDN architecture within the satellite network. This consist on the introduction of an SDN controller that manages the connectivity services across the SDN-based satellite network and makes use of the following interfaces: (1) SBIs for the MC of the interworking and adaptation functions in gateway STs and potentially also user STs; (2) SBIs for the M& C of the BSM bearer services and potentially also some capabilities within the SD lower layers (satellite resources such as a frequency plan, modulation and coding schemes or other satellite specific properties) and (3) NBIs for the M& C of the satellite network flow services from network applications on top of the SDN controller or from external controllers. Candidate SDN data models and protocols for the realization of the SDN-based satellite network architecture also have been studied, namely, ETSI BSM SI-SAP, ONF OF, ONF Microwave Information Model, ONF T-API and IETFYANG models for traffic engineered networks.

Then, based on this integration approach for the realization of E2E SDN-based TE in satellite-terrestrial backhaul networks, the thesis has developed a general model and problem formulation for the analysis of capacity and traffic management strategies in hybrid satellite-terrestrial mobile backhauling networks that rely on SDN for traffic steering at flow level considering a realistic traffic model. This first development allows contributions in the field of sizing a hybrid satellite-terrestrial mobil backhaul network. The presented results give us some fundamental guidelines for deriving design and dimensioning principles to consider resilience as a built-in feature in 5G systems. The main findigs of the capacity assessment are the followings:

- An initial assessment has been reported to estimate the capacity gains that the satellite component could bring when used to support a basic traffic overflow strategy.

- The results obtained have revealed that, besides further motivations supporting the interest of the satellite backhaul (e.g. service extension in hard to reach

areas, increased resilience, temporary deployments), relevant capacity gains can be achieved thanks to the multiplexing gain of sharing some amount of satellite capacity across a number of BS sites and/or A1 traffic aggregation points.

- In particular, it has been shown under a realistic traffic model that if the shared satellite capacity is deployed to accommodate the traffic demand between the 90-th and 99-th percentiles, capacity reduction factors within a range of 5-15 and 4-6 can be achieved with respect to the case that such capacity is provisioned in a dedicated manner at BS or $A1$ levels, respectively.

Subsequently, given the satisfactory results in terms of capacity gains due to the introduction of the satellite component to a terrestrial backhaul network, the research has progressed towards the development of traffic distribution strategies, which allow to further enhancement of the capacity gained. In this regard, this work has focused in the development and assessment of traffic distribution strategies that exploits the dynamically steerable satellite capacity provisioned for resilience purposes to maximize a network utility function under both failure and non-failure conditions in the terrestrial links.

The development of traffic distribution strategies considers the discrepancies in performance that exist between satellite and terrestrial satellite links through the design of utility functions, as well as the incorporation of utility factors applicable to each type of backhaul link and type of traffic. A first development of traffic distribution strategies considers elastic and inelastic traffic (Non-GBR and GBR respectively), and the mathematical development of the utility function according to the number of flows in the network, type of service and number of BSs, as well as the mathematical development and algorithms for their solution. Furthermore, the thesis has numerically assessed by means of simulation the developed traffic distribution strategies. The analysis has considered different resilience schemes, characterized by the ratio $N : M$ of satellite capacity needed to cope with the failure of N terrestrial links in a group of M BSs.

This second development contributes as the first traffic distribution strategies developed for hybrid satellite-terresteial mobile backhaul networks. The main findigs

of the assessmend are the followings:

- Obtained results show how, under failure conditions, the proposed traffic distribution strategy is able to redistribute the satellite resources so that the utility decrease in the BSs affected by terrestrial link failures is minimized.
- Remarkably, when all the terrestrial links are fully operational, it has been shown that the proposed strategy does not leave the satellite capacity unused but exploits it in order to improve the overall network utility.
- The presented results give us some fundamental guidelines for deriving design and dimensioning principles to consider resilience as a built-in feature in 5G systems.

Then, the thesis work deepens in the analysis of the optimality conditions and algorithms to cope with the network utility maximization problem as well as its extension to jointly consider other uses of the 5G backhaul integrated satellite capacity such as fast deployments/special events and the delivery of multicast traffic. The developed TE solution allows managing the use of the satellite capacity provisioned for resilience purposes among a number of BSs so that the overall network utility is maximized under both failure and non-failure conditions in the terrestrial links, considering QoS differentiation and the consideration of multicast traffic.

The developed SDN-basedTE application was formulated that building on a global view of the hybrid terrestrial-satellite network resources exploits a combination of control features and criteria such as (1) E2E path computation with terrestrial or satellite link selection; (2) satellite capacity resource reservations to deal with BSs with no or limited terrestrial link backhaul capacity; (3) different allocation criteria depending on the traffic nature (GBR and non-GBR services, unicast/multicast); (4) admission and rate control to face overload and guarantee resources and minimum (committed) transmission rates per flow and group of flows and (5) utility maximization criteria, where the adequacy of handling specific flows over the terrestrial or satellite component, as well the effect of allocating more or less data rate, are accounted. A detailed performance analysis has been conducted to assess the

behaviour of the proposed SDN-based TE application in multiple and diverse scenarios, including homogeneous and non-homogeneous load situations with BSs that exploit both satellite and terrestrial backhaul capacity, terrestrial link failures in some of the BSs and deployment of a number of mobile TBSs that exclusively rely on the satellite capacity for backhauling. A more traditional overflow strategy has been considered for comparison purposes. As general trends, it has been demonstrated how the proposed SDN-based TE application is able to provide a higher network utility in most of the analysed cases, greatly improving the admission rejection ratio for GBR services and achieving higher fairness in the distribution of delivered data rates for non-GBR flows.

The performance of the proposed TE application has been assessed by means of numerical simulation. The main findigs of the assessmend are the followings:

- Obtained results shows how overall network performance is improved, compared to that of a traditional overflow solution, in terms of:
  - Network utility increase.
  - GBR rejection rate decrease.
  - Non-GBR utility increase.
  - Improved Non-GBR fairness, especially for BSs that temporary face a lack of terrestrial link availability.
  - Under failure conditions, it has been shown that the reservation scheme implemented within the TE application allows keeping fair utility levels in the BSs affected by terrestrial link failures.

Finally, this thesis presents an experimental Proof of Concept (PoC) of a satellite-terrestrial integration solution that built upon Software-Defined Networking (SDN) technologies for the realization of End-to-End Traffic Engineering (E2E TE) in mobile backhauling networks with a satellite component, validating the operation of a E2E TE application able to enforce different traffic routing and path failure restoration policies as well as assessing the performance impact that it has on the mobile network connectivity services. Sustained in the architectural designs proposed by the EU-funded VITAL project, this thesis presented an experimental PoC and validation

of an integration solution based on the use of SDN technologies for the realization of E2E TE applications in hybrid terrestrial-satellite backhaul mobile scenarios.

The implemented testbed contributes in the assessment of the feasibility of the proposed SDN-based integration solution under a practical laboratory setting that combines the use of commercial, experimentation-oriented and emulation equipment and software. Procedures for path computation and restoration have been showcased, assessing the incurred execution times and signalling exchanges incurred in the different steps. In particular, it has been demonstrated that the time incurred in the creation and deletion of flows from the controller is of the order of a few milliseconds, and that the longest delays incurred in the tested procedures that actually have a noticeable impact on service performance are mainly associated with the nature of the hardware used in the experimentation. Moreover, it has been also validated the versatility of using a high-level language like Python and the existing OF libraries in the Ryu controller for programming the TE application in order to enforce different traffic routing and path failure restoration policies. Finally, the testbed has also allowed us to assess the impact on the LTE service performance for TCP and UDP traffic flows when using and switching to/from a satellite link.

8.3 Future Work

Concepts and results presented in this thesis clearly advocate for the need to outfit next-generation satellite networks with a set of control and management functions and interfaces (API and/or network protocols) compatible with the mainstream SDN architectures and technologies being adopted in 5G in order to realize a full E2E networking concept where a combined satellite-terrestrial network service can be deployed and operated in a flexible and consistent manner. While the SDN-based solution explored in this thesis has exclusively focused on the ground segment of satellite communications systems, dynamic interactions with flexible satellite payloads for more efficient resource management across the whole satellite communications chain should also be explored.

This thesis has set the stage for the deployment of innovative SDN applications, targeting, for example, enhanced resource efficiency, efficient and fast protection and

restoration, and automation of network planning and operation in network infrastructures spanning terrestrial and satellite resources.

Finally, while the results achieved in the PoC presented in this thesis are clearly supportive of the potential and feasibility of applying SDN technologies for improved terrestrial-satellite integration, further work is still necessary to demonstrate additional control features enabled by OF such as per-flow metering, rate-limiting, packet duplication for multicast, etc., which could be leveraged for the implementation of more complex traffic distribution and QoS policing operations.

Bibliography

[1] NGMN Alliance, *5G White Paper*, February 2015, available at https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf.

[2] NetWorld2020 – SatCom Working Group, *The Role of Satellites in 5G*, Version 5, 31st July 2014. Available at https://www.networld2020.eu/wp-content/uploads/2014/02/SatCom-in-5G_v5.pdf.

[3] Osseiran A. et al., *Scenarios for 5G Mobile and Wireless Communications: The Vision of the METIS Project*, IEEE Communications Magazine, May 2014.

[4] 5G Vision, *The 5G Infrastructure Public Private Partnership: the next generation of communication networks and services*, available at https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf.

[5] Sacchi, C.; Bhasin, K.; Kadowaki, N.; Vong, F., *Toward the "space 2.0" Era [Guest Editorial]*, Communications Magazine, IEEE , vol.53, no.3, pp.16,17, March 2015.

[6] Fenech, H.; Amos, S.; Tomatis, A.; Soumpholphakdy, V., *High throughput satellite systems: An analytical approach*, in Aerospace and Electronic Systems, IEEE Transactions , vol.51, no.1, pp.192-202, Jan. 2015.

[7] ARTES programme, *ESA announces dedicated support for the development of megaconstellations*, available at https://artes.esa.int/news/esa-announces-dedicated-support-development-megaconstellations. Last updated at 02 February 2016.

[8] 3GPP RP-171450, *Study on NR to support Non-Terrestrial Networks*, 3GPP TSG RAN WG1 Meeting 88bis, West Palm Beach, USA, 5th–9th June 2017.

[9] 3GPP TS 22.261 V0.1.1 (2016-08), *Service requirements for next generation new services and markets; Stage 1 (Release 15)*, August 2016.

[10] 3GPP a Global Initiative,The Mobile Broadband Standard, *Satellite components for the 5G system*, available at http://www.3gpp.org/news-events/3gpp-news/1933-sat_ntn. Last updated at January 4, 2018.

[11] Gomes T. et al., *A survey of strategies for communication networks to protect against large-scale natural disasters*, 8th International Workshop on Resilient Networks Design and Modeling (RNDM). 13-15 Sept. 2016; Halmstad, Sweden.

[12] Watts S., Glenn O., *5G resilient backhauling using integrated satellite networks*, Advanced Satellite Multimedia System Conference, September 2014.

[13] Breiling M., Zia W., Sanchez de la Fuente Y., Mignone V., Milanesio D., Fan Y., Guta M., *LTE Backhauling Over MEO-Satellite*, Advanced Satellite Multimedia Systems Conference, pp. 174-181, IEEE September 2014.

[14] Bertaux L. et al., *Software Defined Networking and Virtualization for Broadband Satellite Networks*, IEEE Communications Magazine, pp. 54 – 60, March 2015.

[15] Toufik A. et al., *Software-defined satellite cloud RAN*, International Journal of Satellite Communications and Networking, Volume36, Issue1, pp. 108-133, February 2017.

[16] H2020 SaT5G Project, *Satellite and Terrestrial Network for 5G, Start date June 2017.*, Project website at http://sat5g-project.eu/

[17] ARTES Telecommunications & Integrated Applications, *Satellite for 5G*, Available at http://www.esa.int/Our_Activities/Telecommunications_Integrated_Application s/Satellite_for_5G. Last update at 21 June 2017.

[18] H2020 VITAL Research Project, 2015. *Satellite for 5G*, Available at http://www.ictvital.eu/. Last Accessed 1 September 2017.

[19] R. Ferrus, H. Koumaras, O. Sallent, G. Agapiou, T. Rasheed, M.-A. Kourtis, C. Boustie, P. Gelard, T. Ahmed., *SDN/NFV-enabled satellite communications networks: Opportunities, scenarios and challenges*, Elsevier, Physical Communication, Volume 18, Part 2, Pages 95-112. March 2016.

[20] Giambene G., Kota S. and Pillai P., *Satellite-5G Integration: A Network Perspective*. in IEEE Network, vol. 32, no. 5, pp. 25-31, September/October 2018.

[21] Xu S., Wang X. and Huang M., *Software-Defined Next-Generation Satellite Networks: Architecture, Challenges, and Solutions*. in IEEE Access, vol. 6, pp. 4027-4041, 2018.

[22] Li T., Zhou H., Luo H., and Yu S., *SERvICE: A Software Defined Framework for Integrated Space-Terrestrial Satellite Communication*. in IEEE Transactions on Mobile Computing, vol. 17, no. 3, pp. 703-716, 1 March 2018.

[23] Bojic D, Sasaki E, Cvijetic N, Ting W, Kuno J, Lessmann J, Schmid S, Ishii H, Nakamura S., *Advanced wireless and optical technologies for small-cell mobile backhaul with dynamic software-defined management*. IEEE Communications Magazine 51(9):86-93.

[24] NGMN Alliance,. *Guidelines for LTE backhaul traffic estimation*, White paper.2011. Available at `https://www.ngmn.org/uploads/media/NGMN_Whitepaper_Guideline_for_LTE_Backhaul_Traffic_Estimation.pdf`.

[25] Zaki Y., Hauth S., Wallmeier E., Gorg C., *Intelligent Traffic Enforcement for LTE Backhaul*, IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). 2013.

[26] Chen-Yui Yang C., Ketcham M., Lu D. and Kinsey D., *Performance Monitoring with Predictive QoS Analysis of LTE Backhaul*, IEEE International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). 2011.

[27] Raza H., *A Brief Survey of Radio Access Network Backhaul Evolution: part II*, IEEE Communications Magazine. Volume:51, Issue: 5. Pp. 170-177. May 2013.

[28] ETSI TR 103 124 V1.1.1 (2013-07), *Satellite Earth Stations and Systems; Combined Satellite and Terrestrial Networks scenarios*.

[29] Viasat,. *Guidelines for LTE backhaul traffic estimation*. Available at https://www.viasat.com/news/going-global. Last accessed 4 December 2018.

[30] Whitepaper,. *The Intelsat EpicNG Platform: High Throughput, High Performance to Support Next-Generation Requirements.*. Available at http://www.intelsat.com/wpcontent/uploads/2013/03/Epic$_W hite_P aper.pdf$.

[31] Ericsson Mobility Report, June 2014.

[32] Casoni M., Gracia C., Klapez M., Patriciello N., Amditos A., Sdongos E., *Integration of Satellite and LTE for Disaster Recovery*. IEEE Communications Magazine, March 2015, pp. 47-53.

[33] Whitepaper, *Extending 3G and 4G Coverage to Remote and Rural Areas Solving The Backhaul Conundrum*. iDirect.

[34] Whitepaper, *Satellite Backhaul for Rural Small Cells*. informa telecoms & media. 2012.

[35] Whitepaper, *The case for rural data services and satellite backhaul*. informa telecoms & media. 2012.

[36] Whitepaper, *Backhaul for rural and remote small cells*. Small cell forum, release 5, March 2015.

[37] 3GPP TS 23.214, *Architecture enhancements for control and user plane separation of EPC nodes*. Release 14. June 2016.

[38] 3GPP TS 23.501, *System Architecture for the 5G System*. Release 15, December 2016.

[39] T. Rossi, M. De Sanctis, E. Cianca, C. Fragale, M. Ruggieri and H. Fenech, *Future Space-based Communications Infrastructures based on High Throughput Satellites and Software Defined Networking*. IEEE International Symposium on Systems Engineering (ISSE), 2015.

[40] ONF TR-521, *SDN Architecture*. Issue 1.1, February 2016.

[41] E. Haleplidis and K. Pentikousis (Editors), *Software-Defined Networking (SDN): Layers and Architecture Terminology*. IRTF RFC 7426, January 2015.

[42] ONF TS-025, *OpenFlow Switch Specification*. Version 1.5.1, March 2015.

[43] ONF TR-535, *ONF SDN Evolution*. Version 1.0, September 2016.

[44] ONF TR-513, *Common Information Model (CIM) Overview 1.2*. September 2016.

[45] ONF TR-512, *Core Information Model (CoreModel) 1.2*. September 2016.

[46] ONF TR-522, *SDN Architecture for Transport Networks*. March 2016.

[47] ONF TR-527, *Functional Requirements for Transport API*. June 2016.

[48] ONFTR-532, *Microwave Information Model*. Version 1.0, December 2016.

[49] ONF TR-523, *Intent Definition Principles*. October 2016.

[50] M. Bjorklund (Editor), *TheYANG 1.1 Data Modeling Language*. IETF RFC 7950, August 2016.

[51] M. Bjorklund, *YANG – A Data Modeling Language for the Network Configuration Protocol (NETCONF)*. RFC 6020, October 2010.

[52] A. Bierman, Bjorklund M. andWatsen K., *RESTCONF Protocol*. IETF RFC 8040, January 2017.

[53] Y. Lee, X. Zhang,D. Ceccarrelli, B.Y.Yoon,O. G. de Dios and J.Y. Shin, *Applicability of YANG models for Abstraction and Control of Traffic Engineered Networks*. June 2017, draft-zhang-teas-actn-yang-05.

[54] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka and T. Turletti, A., *Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks*. IEEE Communications Surveys & Tutorials, Third Quarter 2014.

[55] C. Janz, L. Ong, K. Sethuraman and V. Shukla, *Emerging Transport SDN Architecture and Use Cases*. IEEE Communications Magazine, October 2016.

[56] ETSI TR 101 984, *Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM); Services and Architectures*. December 2007.

[57] ETSI TR 102 187, *Overview of BSM Families*. May 2003.

[58] ETSI TS 102 357, *CommonAir Interface Specification; Satellite Independent Service Access Point SI-SAP*. May 2005.

[59] ETSI TR 103 444, *Guide to Satellite Independent Service Access Point (SISAP) Use*. December 2016.

[60] ETSI TS 102 462, *QoS Functional Architecture*. June 2015.

[61] ETSI TS 102 295, *BSM Traffic Classes*. February 2004.

[62] VITAL Deliverable D4.5, *Final network resource management framework and performance assessment of algorithmic solutions*. R. Ferrus (Editor), August 2018. Available at `http://www.ict-vital.eu/documents/deliverables`. Last accessed in December 2018.

[63] ETSI TS 103 275, *Satellite Independent Service Access Point (SI-SAP) interface: Services*. May 2015.

[64] S. Abdellatif, P. Berthou, P. Gelard, T. Plesse and S. El-Yousfi, *Exposing an Openflow Switch Abstraction of the Satellite Segment to Virtual Network Operators*. 2016 IEEE 83rdVehicular Technology Conference (VTC Spring), Nanjing, 2016.

[65] D. Bogdanovic, B. Claise and C. Moberg, *YANG Module Classification*. IETF Internet Draft, October 2016, draft-ietf-netmod-yangmodelclassification.

[66] I. Busi (Editor), *Transport Northbound Interface Use Cases*. July 2017, draft-tnbidt-ccamp-transport-nbi-use-cases-02.

[67] W. Liu and A. Farrel, *Service Models Explained*. IETF Internet Draft, June 2017, draft-ietf-opsawg-service-model-explained-01.

[68] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell and J. McManus, *Requirements for Traffic Engineering over MPLS*. IETF RFC 2702, September 1999.

[69] M. Conran, *OpenFlow Traffic Engineering*. September 2015. Available at `http://network-insight.net/2015/09/openflow-traffic-engineering/`. Last Accessed 1 July 2017.

[70] D. Bojic, E. Sasaki, N. Cvijetic, et al., *Advanced Wireless and Optical Technologies for Small-cell Mobile Backhaul with Dynamic Software-defined Management*. IEEE Communications Magazine, pp. 86–93, September 2013.

[71] M. R. Sama, L. M. Contreras, J. Kaippallimalil, I. Akiyoshi, H. Qian and H. Ni, *Software-defined Control of the Virtualized Mobile Packet Core*. IEEE Communications, pp. 107–115, February 2015.

[72] VITAL Deliverable D4.2, *Network resource management framework and initial performance assessment of algorithmic solutions*. R. Ferrus (Editor), August 2018. Available at `http://www.ict-vital.eu/documents/deliverables`. Last accessed in December 2018.

[73] N. L. M. van Adrichem, C. Doerr and F. A. Kuipers, *OpenNetMon: Network Monitoring in OpenFlow Software-Defined Networks*. IEEE Network Operations and Management Symposium (NOMS), Krakow, 2014.

[74] Aricent White Paper, *Demystifying Routing Services in Software-defined Networking*. November 2014. Available at `http://www.aricent.com/sites/default/files/pdfs/Aricent-Demystifying-Routing-Services-SDNWhitepaper.pdf`.

[75] R. Pujar and I. Camelo, *Path Protection and Failover Strategies in SDN Networks*. Inocybe Technologies, Open Networking Summit, March 2016.

[76] U. Gotzner, A. Gamst, and R. Rathgeber, *Spatial traffic distribution in cellular networks*. in Proc. IEEE VTC, 1998.

[77] J. Reades, F. Calabrese, C. Ratti, *Eigenplaces: analysing cities using the space - time structure of the mobile phone network*. Environment and Planning B: Planning and Design. vol. 36(5), pp. 824 - 836, 2009.

[78] Lee, Zhou, *Spatial Modelling of the Traffic Density in Cellular Networks*. IEEE Wireless Communications, February 2014.

[79] D. Lee, S. Zhou, and Z Niu, *Spatial Modeling of Scalable Spatially-Correlated Log-Normal Distributed Traffic Inhomogeneity and Energy-Efficient Network Planning*.

IEEE Wireless Communications and Networking Conference (WCNC) IEEE, 2013.

[80] James Allen, Franck Chevalier, Burcu Bora, *Mobile backhaul market: Phase 1 report*. Analysis Mason, February 2014.

[81] N. Cassiau and D. Ktenas, *Satellite Multicast for Relieving Terrestrial eMBMS:System-Level Study*. 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall), Boston, MA, 2015, pp. 1-5.

[82] E. H. K. Wu and Chao-hsu Chang, *Adaptive multicast routing for satellite-terrestrial network,*. Global Telecomm. Conference, 2001. GLOBECOM '01. IEEE, San Antonio, TX, 2001, pp. 1440-1444 vol.3.

[83] N. Chuberre1, O. Courseille1, P. Laine2, L. Roullet2, T. Quignon1 and M. Tatard, *Hybrid satellite and terrestrial infrastructure for mobile broadcast services delivery: An outlook to the 'Unlimited Mobile TV' system performance*. International journal of satellite communications int. J. Commun. Syst. Network 2008; 26:405–426 Published online in Wiley InterScience `www.interscience.wiley.com`. DOI: 10.1002/sat.910.

[84] Laxman H. Sahasrabuddhe and Biswanath Mukherjee, *Multicast Routing Algorithms and Protocols: A Tutorial*. IEEE Net., Jan 2000.

[85] Zhizhong Yin, Long Zhang, Xianwei Zhou, Peng Xu, Yu Deng, *QoS Guaranteed Secure Multicast Routing Protocol for Satellite IP Networks Using Hierarchical Architecture*. Int. J. Communications, Network and System Sciences, 2010.

[86] Eylem E., Ian F. Akyildiz, and Bender M., *A Multicast Routing Algorithm for LEO Satellite IP Networks*. IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 10, no. 2, April 2002.

[87] De-Nian Yang, Wanjiun Liao, *On multicast routing using rectilinear Steiner trees for LEO satellite networks*. IEEE Comm. Society, Globecom 2004.

[88] C. A. Grazia et al., *Integration between terrestrial and satellite networks: the PPDR-TC vision*. 2014 IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Larnaca, 2014.

[89] E. H. Fazli, M. Werner, N. Courville, M. Berioli and V. Boussemart, *Integrated GSM/WiFi Backhauling over Satellite: Flexible Solution for Emergency Communications*. Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE, Singapore, 2008.

[90] SANSA (Shared Access Terrestrial-Satellite Backhaul Network enabled by Smart Antennas) research project, HORIZON 2020 Framework Programme. Website: `http://www.sansa-h2020.eu/`.

[91] S.Shenker, *Fundamental Design Issues for the Future Internet*. IEEEJ-SAC, vol.13, no.7, pp.1176-1188, Sep.1995.

[92] Zhimei Jiang, Ye Ge, Ye Li, *Max-utility wireless resource management for best-effort traffic*. IEEE Trans. Wireless Comm, vol.4, no.4, pp.100-111, Jan.2005.

[93] E. Kelly, *Charging and rate control for elastic traffic*. European Trans. Telecomm, vo1.8, pp.33-37, 1997.

[94] S. Burera, A. N. Letchfordb, *Non-convex mixed-integer nonlinear programming: A survey*. Surveys in Operations Research and Management Science, Volume 17, Issue 2, July 2012, Pages 97-106.

[95] T. Harks, *Utility Proportional Fair Bandwidth Allocation: An Optimization Oriented Approach*. in QoS-IP, pp. 61-74, 2005.

[96] ESA news: `https //artes.esa.int/news/newtec-introduces-industrysfirst-dvb-s2x-vsat -modem`. Last access on 18 November 2016.

[97] *Video as a Basic Service of LTE Networks: Mobile vMOS Defining Network Requirements* `http://www.huawei.com/minisite/4-5g/en/industryjsdc-j.html`. Huawei. 2015.

[98] OpenSignal. *The State of LTE*. February 2016. Available at `http://opensignal.com/reports/2016/02/state-of-lte-q4-2015/`.

[99] *OpenFlow Traffic Engineering*. Available at `http://network-insight.net/2015/09/openflow-traffic-engineering/`. Last access on 5 December 2018.

[100] B. Heller et al., *ElasticTree: Saving energy in data center networks*. in Proc. 7th USENIX Conf. Netw. Syst. Design Implement., 2010, p. 17.

[101] M. Al-Fares, S. Radhakrishnan, B. Raghavan, N. Huang, and A. Vahdat, *Hedera: Dynamic flow scheduling for data center networks*. in Proc. 7th USENIX Conf. Netw. Syst. Design Implement., 2010, p. 19.

[102] R. Wang, D. Butnariu, and J. Rexford, *OpenFlow-based server load balancing gone wild*. in Proc. 11th USENIX Conf. Hot Topics Manage. Internet Cloud Enterprise Netw.Services, 2011, p. 12.

[103] P. Xiong and H. Hacigu mu s, *Pronto: A software-defined networking based system for performance management of analytical queries on distributed data stores*. PVLDB, vol. 7, no. 13, pp. 1661–1664, 2014.

[104] K. Jeong, J. Kim, and Y.-T. Kim, *QoS-aware network operating system for software defined networking with generalized OpenFlows*. in Proc. IEEE Netw. Oper. Manage. Symp., Apr. 2012, pp. 1167–1174.

[105] S. Sharma et al., *Implementing quality of service for the software defined networking enabled future internet*. in Proc. 3rd Eur. Workshop Softw. Defined Netw., 2014, pp. 49–54.

[106] W. Kim et al., *Automated and scalable QoS control for network convergence*. in Proc. Internet Netw. Manage. Conf. Res. Enterprise Netw., 2010.

[107] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky and S. Uhlig, *Software-Defined Networking: A Comprehensive Survey*. in Proceedings of the IEEE, vol. 103, no. 1, pp. 14-76, Jan. 2015.

[108] X.-N. Nguyen, D. Saucez, C. Barakat, and T. Turletti, *Optimizing rules placement in OpenFlow networks: Trading routing for better efficiency*. in Proc. 3rd Workshop Hot Topics Softw. Defined Netw., 2014, pp. 127–132.

[109] A. Schwabe and H. Karl, *Using MAC addresses as efficient routing labels in data centers*. in Proc. 3rd Workshop Hot Topics Softw. Defined Netw., 2014, pp. 115–120.

[110] I. F. Akyildiz, A. Lee, P. Wang, M. Luo, and W. Chou, *A roadmap for traffic engineering in SDN-OpenFlow networks*. Comput. Netw., vol. 71, pp. 1–30, Oct. 2014.

[111] Prashanth Hande, Shengyu Zhang, and Mung Chiang, *Distributed Rate Allocation for Inelastic Flows*. IEEE/ACM Transactions on Networking, vol, 15, No. 6 , December 2007.

[112] Dah Ming Chiu, Adrian Sai-Wah Tam, *Fairness of Traffic Controls for Inelastic Flows In the Internet*. 2006.

[113] A. Ouorou, P. Mahey, and J.-Ph. Vial, *A survey of algorithms for convex multicommodity flow problems*. Manage. Sci., vol. 46, pp. 126–147, Jan. 2000.

[114] Jang-Won Lee, Ravi R. Mazumdar, Ness B. Shroff, *Non-Convex Optimization and Rate Control for Multi-class Services in the Internet*. IEEE/ACM Transactions on Networking, vol 13, No. 4, August 2005.

[115] Ferrús R., Sallent O., Ahmed T., Fedrizzi R. *Towards SDN/NFV-enabled satellite ground segment systems: End-to-End Traffic Engineering Use Case*. IEEE International Conference on Communications Workshops (ICC Workshops). 21- 25 May 2017; Paris.

[116] ONF Open Networking Foundation *OpenFlow Switch Specification*. Version 1.3.0 (Wire Protocol 0x04). 25 June 2012. Available at `https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-spec-v1.3.0.pdf`.

[117] AMARI LTE 100. Available at `https://www.amarisoft.com`. Last accessed 13 October 2018.

[118] OpenSAND. Available at `http://opensand.org`. Last accessed 15 October 2018.

[119] Ryu. Available at `https://osrg.github.io/ryu/index.html`. Last accessed 15 October 2018.

[120] Northbound Networks. Zodiac FX Product. Available at https://northboundnetworks.com/collections/zodiac-fx.. Last accessed 15 October 2018.

[121] Miller C. *How and why commercial high-capacity satellites offer superior performance and survivability in the future space threat continuum*. 32nd Space Symposium, Technical Track. April 11-12, 2016. Colorado Springs, Colorado, USA.

[122] Whitepaper. *Why Latency Matters to Mobile Backhaul - O3b Networks*. O3b Networks. 2013.

[123] Iperf. *The ultimate speed test tool for TCP, UDP and SCTP* Available at https://iperf.fr/. Last accessed 15 October 2018.

Appendix A

Network Utility Maximization Simulator

It has been demonstrated that a proportional fair resource allocation ensures a max-min fairness for all users sharing a single path in the network [94]. Then, a max-min rate allocation is equivalent to the maximization of the utility at each link for UN traffic. This same concept can be extended to the global utility of the network when there is no distinction between using satellite or terrestrial capacity, which in our case is strictly valid for factor $p^{UN} = 1$.

This can be clearly observed when developing the components of the utility function to be maximized, presented in chapter 5 (eq. 5.2), considering only UN traffic. This is:

Max:

$$\sum_{s \in \{UN\}} U^S(r_{s,m,i}, x_{s,m,i}) \tag{A.1}$$

s.t. Eq. Eq. 5.11-Eq. 5.14

Considering $p^{UN} = 1$ and developing the components of the previous equation, we have that the network utility is the sum of all UN utility functions. This can be rewritten as follows:

$$U(x) = log(r_1) + log(r_2) + \cdots + log(r_I) = log(r_1 \cdot r_2 \cdots r_I) \tag{A.2}$$

where I is the total number of flows in the network. From eq. Eq. A.2, is observed that the maximum network utility is obtained when the product of bit rates

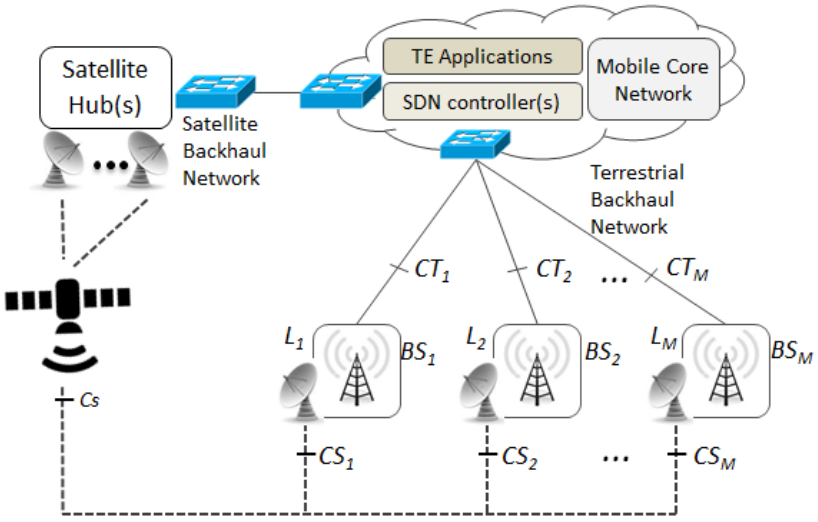

FIGURE A.1: Network Model.

is maximized. Therefore, the product is maximized when all the bit rates are equal and as high as possible, according to the capacity restrictions.

This resource distribution is made by an heuristic "Network Utility Maximization (NUM) Simulator". The heuristic structure is presented as follows:

We utilize the nomenclature presented in Figure A.1.

where:

M : Number of BS's that share satellite capacity.

L_i : Number of UN flows of the i-th BS.

CT_i : Terrestrial capacity of the i-th BS available for handling UN traffic.

C_S : Aggregated satellite capacity available for handling UN traffic.

TR_i : Terrestrial Bit rate assigned per UN flow in the i-th BS.

SR_i : Satellite Bit rate assigned per UN flow in the i-th BS.

CS_i : Satellite capacity assigned to the i-th BS.

With the assumption of the following considerations:

1. The capacity available for UN traffic in a given link, is shared equally among all the UN flows handled by the link.

2. The algorithm considers only UN flows.

The algorithm consists in two main parts, captured in 2 main Matlab functions as follows:

Function 1: **[CSi] = Cap_ Sat_ UN(t_ index,config,BS);**

This first function finds the satellite capacity assigned to each BS (CS_i), in such a way that all satellite and terrestrial UN connections have the same bit rate or, if this is not possible, the one closest to each other. Also, the assigned bit rate must be the highest possible, subject to capacity restrictions. To achieve this, a first approximation is the obtaining of the capacity assigned to each BS in proportion to the number of flows served, as shown in the following expression:

$$\frac{CT_1 + CS_1}{L_1} + \frac{CT_2 + CS_2}{L_2} = \cdots = \frac{CT_M + CS_M}{L_M} \tag{A.3}$$

The eq. A.3 indicates that the total amount of capacity (terrestrial and satellite capacity) assigned to each BS, divided by the number of flows handled by each BS must be the same for all BS's. Since we know the terrestrial capacity and the number of flows per BS, we organize the expression as follows:

$$FT_1 + \frac{CS_1}{L_1} + FT_2 + \frac{CS_2}{L_2} = \cdots = FT_M + \frac{CS_M}{L_M} \tag{A.4}$$

where $FT_i = C_i/L_i$ represents the terrestrial capacity per flow when there is no satellite capacity assigned to the i-th BS. Therefore, so that Eq. A.4 is fulfilled, each term $FT_i + CS_i/L_i$ must be sufficient to at least reach the highest FT_i value (FT_{higher}). In case the satellite capacity is not enough to comply with the Eq. A.4 for all BSs, the BSs with the highest FT_i are discarded consecutively until the expression is fulfilled ($CS_i = 0$ for discarded BSs). In case the expression is fulfilled, the distribution of satellite resources is carried out in two parts; a first part of the satellite capacity ($CS_{i-first}$) is distributed so that the expression is fulfilled and; the second part consists in distributing the remaining satellite capacity ($C_{remaining}$) to each BS, proportionally to their number of flows ($CS_{i-second}$).

To do this, the implemented heuristic algorithm follows the following steps:

1. The FT_i values are calculated.

2. The higher FT_i value is obtained (FT_{higher}).

3. The $CS_{i-first}$ for all BSs needed to equal at least the FT_{higher} value is obtained with the following expression:

$$CT_{higher} = CT_i + \frac{CS_{i-first}}{L_i} \qquad \forall i \tag{A.5}$$

4. In case the satellite capacity is not enough to comply with the previous expression for all BS, the BSs with PT_i with higher value (PT_{higher}) are discarded and assigned $CS_{i-first} = 0$, consecutively until the expression is fulfilled. For $CS_{i-first}$, reorganizing the previous expression we have:

$$CS_{i-first} = (FT_{higher} - FT_i) \cdot L_i, \qquad \forall i \tag{A.6}$$

subject to:

$$\sum_{i=1}^{M} CS_{i-first} \leq C^S \tag{A.7}$$

5. The remaining satellite capacity is calculated by the following expression:

$$C_{remaining} = C^S - \sum_{i=1}^{M} CS_{i-first} \leq C^S \tag{A.8}$$

6. The remaining satellite capacity is distributed among the BSs in proportion to the number of flows by:

$$CS_{i-second} = \frac{L_i}{C^S - \sum_{i=1}^{M} L_i} C_{remaining} \tag{A.9}$$

7. Finally, CS_i is calculated as follows:

$$CS_i = CS_{i-first} + CS_{i-second}, \qquad \forall i \tag{A.10}$$

In this way, the satellite capacity is distributed in such a way that all BSs have a backhaul capacity (terrestrial + satellite) proportional to the number of flows. It is clear that this does not indicate that all the flows of each BS will reach the same bit rate, because each BS has two backhaul links and the number of connections will have to be divided between them, however, this is a first approach. Subsequently, once the satellite capacity value assigned to each BS is defined, the algorithm defines the number of flows to be served by each link through some iterations, in such a way that the bit rate assigned to the flows is the most similar possible in order to obtain the highest network utility. This part is made by the next function:

Function 2: **[Tasa_ terr,Tasa_ sat,Con_ terr,Con_ sat,Utility]= Optimum_ distribution (CSi,t_ index,config,BS)**

The function obtain the number of flows to be sent per link type, as well as the terrestrial and satellite capacity assigned to each UN flow (TR_i , SR_i) in such a way that network utility is maximized.

The function takes the value CS_i, and defines the number of terrestrial and satellite flows at each each BS in order to obtain the UN bit rates (TR_i and SR_i) as close as possible. For $p^{UN} < 1$ the algorithm makes some iterations increasing and decreasing the number of flows at each BSs to find the number of flows to be handled at each link such that the utility is maximized at each BS.

Appendix B

TE-Application Simulator

The input variables of the simulator are presented below which we can divide into the following groups:

Simulation parameters:

Variables:	**Description:**
-config.simulation_ duration	-Total duration of the simulation (sec).
-config.time_ step	-Simulation time step duration (sec). It is the discrete time value with which the total simulation time is divided, presenting in each time step the results of the simulation.

Network link capacities:

Variables:	**Description:**
-config.num_ cells	-Total number of cells (BS)
-config.BS_ TERRESTRIAL_ CAPACITY	-Terrestrial link capacity per BS (Mbps)
-config.BS_ SATELLITE_ CAPACITY	-Maximum satellite link capacity per BS (Mbps)
-config.GLOBAL_ SATELLITE_ CAPACITY_ RATIO	-Satellite aggregated capacity (ratio of total terrestrial capacity)
-BS(n).capacity_ Terrestrial	-Number of BSs without terrestrial capacity availability.

Services and utility parameters:

Variables:	**Description:**
-config.R1UG	-Rate R1 for UG traffic.
-config.R2UG	-Rate R2 for UG traffic
-config.R1MG	-Rate R1 for MG traffic
-config.R1UN	-Rate R1 for UN traffic
-config.alpha	-Alpha factor (α)
-config.p	-p factor for UG, MG and UN traffic.

Traffic parameters:

Variables:	**Description:**
-BS(n).lambda(1) (vector)	-UG flow generation rate ()
-BS(n).lambda(2) (vector)	-UN flow generation rate ()
-config.traffic_ params.duration(1) (vector)	-Average UG flow duration (sec)
-config.traffic_ params.duration(2) (vector)	-Average UN flow duration (sec)

Network link capacities:

Variables:	**Description:**
-config.Reserva_ Inicial	-satellite capacity value as initial reservation (Percentage of global satellite capacity)
-config.valor_ maximo_ de_ reserva=xx (Mbps)	-Satellite capacity as reservation (Mbps)
-config.umbral_ maximo_ de_ reserva=0.95 (%)	-Maximum satellite capacity reservation (% of global satellite capacity)
-config.time_ window_ utilisation_ averaging_ samples	-Number of time steps to calcultate the utilization for reservation purposes.

The simulator yields results of two types; Numerical and graphic. The numerical results are described below, which we can divide into the following groups:

Global Network Metrics (considering the total number of BSs):

Variables:	**Description:**
-Number of UG flows	-Number of arrived UG flows
-Blocking	-Percentage of blocking UG flows
-Average Bit Rate (TERRESTRIAL)	-Average bit rate for terrestrial UG flows (Mbps)
-Average Bit Rate (SATELITE)	-Average bit rate for satellite UG flows (Mbps)
-Average Bit Rate (BOTH LINKS)	-Average bit rate for UG flows (both kind of links)
-Average Utility per UG service	
-Number of UN flows	-Number of arrived UN flows
-Blocking	-Percentage of blocking UN flows
-Average Bit Rate (TERRESTRIAL)	-Average bit rate for terrestrial UN flows (Mbps)
-Average Bit Rate (SATELITE)	-Average bit rate for satellite UN flows (Mbps)
-Average Bit Rate (BOTH LINKS)	-Average bit rate for UN flows (both kind of links)
-Average Utility per UN service	
-Average Satellite Throughput	-Aggregated satellite capacity occupation (%).
-Average Terrestrial Throughput	-Terrestrial capacity occupation (%).
-Average Link Reasignations	
-Utility	-Average Global Utility
-Average Utility per Flow	-Average Utility per service (considering both UG and UN Flows)
-UN Standard Deviation	-UN bit rates standard deviation ()

BSs with terrestrial capacity (considering the BSs with terrestrial capacity availability):

Variables:	**Description:**
-Number of UG flows	-Number of arrived UG Flows.
-Blocking	-Percentage of blocking UG Flows
-Average Bit Rate	-Average Bit Rate for terrestrial UG Flows (Mbps)
-Average Utility per UG service	-Rate R1 for UN traffic
-Number of UN flows	-Number of arrived UN Flows
-Blocking	-Percentage of blocking UN Flows.
-Average Bit Rate	-Average Bit Rate for terrestrial UN Flows (Mbps)
-Average Utility per UN service	
-Average Terrestrial Throughput	-Aggregated Terrrestrial capacity occupation (%).
-Utility	-Average Global Utility
-Average Utility per Flow	-Average Utility per service (considering both UG and UN Flows)

BSs without terrestrial capacity (considering the BSs with no terrestrial capacity availability):

Variables:	**Description:**
-Number of UG flows	-Number of arrived UG Flows
-Blocking	-Percentage of blocking UG Flows
-Average Bit Rate	-Average Bit Rate for UG Flows
-Average Utility per UG service	
-Number of UN flows	-Number of arrived UN Flows
-Blocking	-Percentage of blocking UN Flows
-Average Bit Rate	-Average Bit Rate for UN Flows
-Average Utility per UN service	
-Average Satellite Throughput	-Aggregated Satellite capacity occupation (%).
-Utility	-Average Utility
-Average Utility per Flow	-Average Utility per service

The results presented graphically are diverse, they show the behavior of different metrics during the simulation time for better analysis. The Figures B.1-B.3 present some of the results as an example below:

The general structure of the Network Utility Maximization Simulator is depicted in the following diagram.

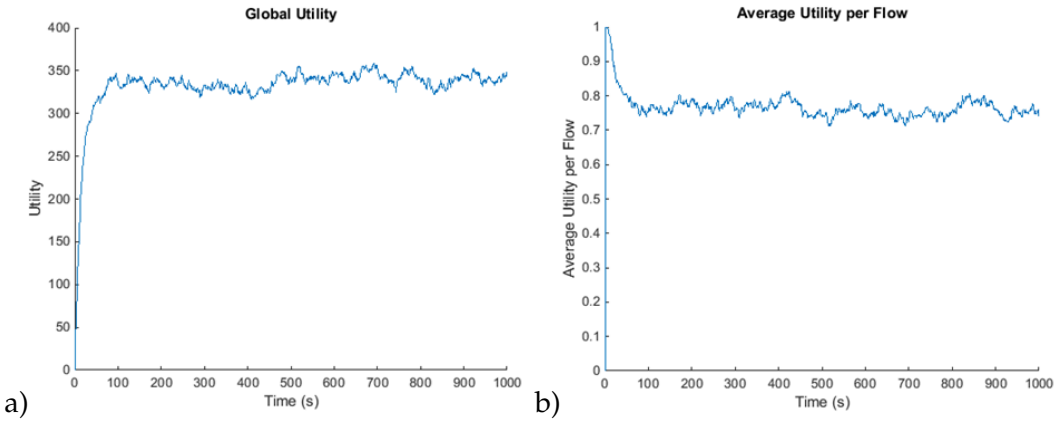

FIGURE B.1: Simulator Results Examples of Time Evolution; (a) Global Utility and (b) Average Utility per Flow.

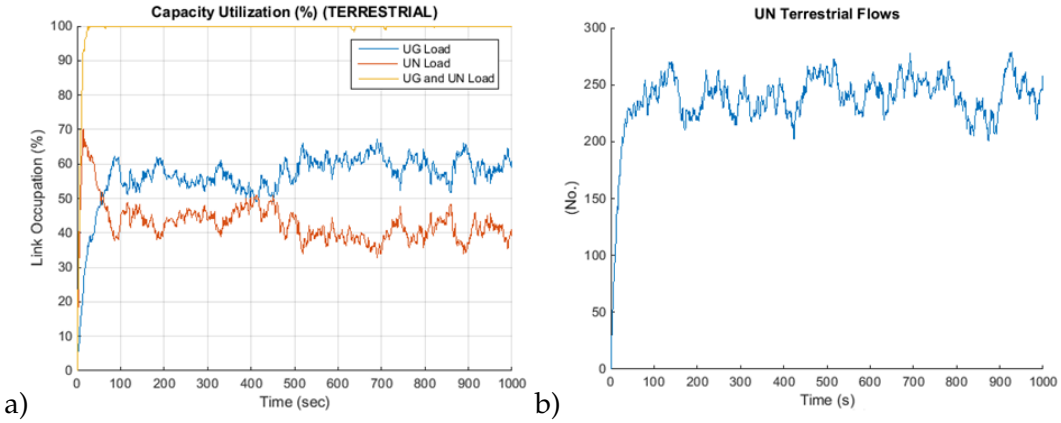

FIGURE B.2: Simulator Results Examples of Time Evolution; (a) Link Occupation and (b) Number of UN Terrestrial Flows.

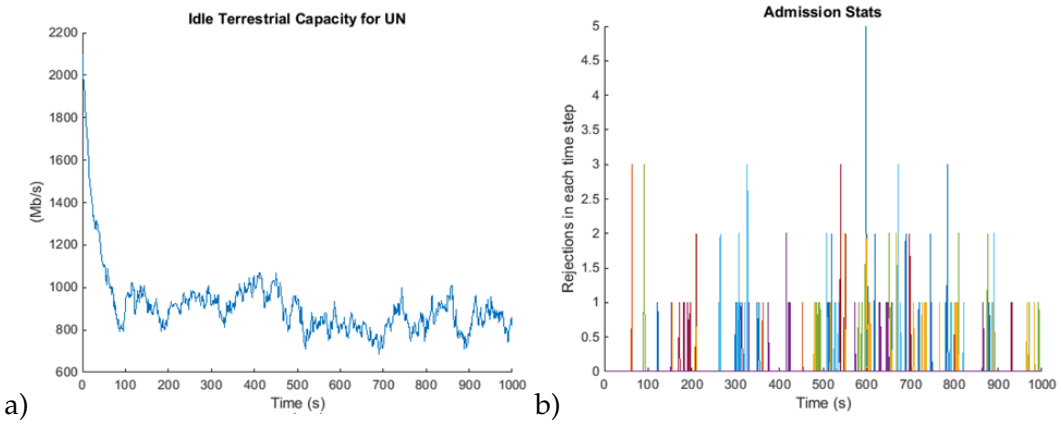

FIGURE B.3: Simulator Results Examples of Idle Terrestrial Capacity; (a) UN Flows and (b) Number of Rejections per Time Step.

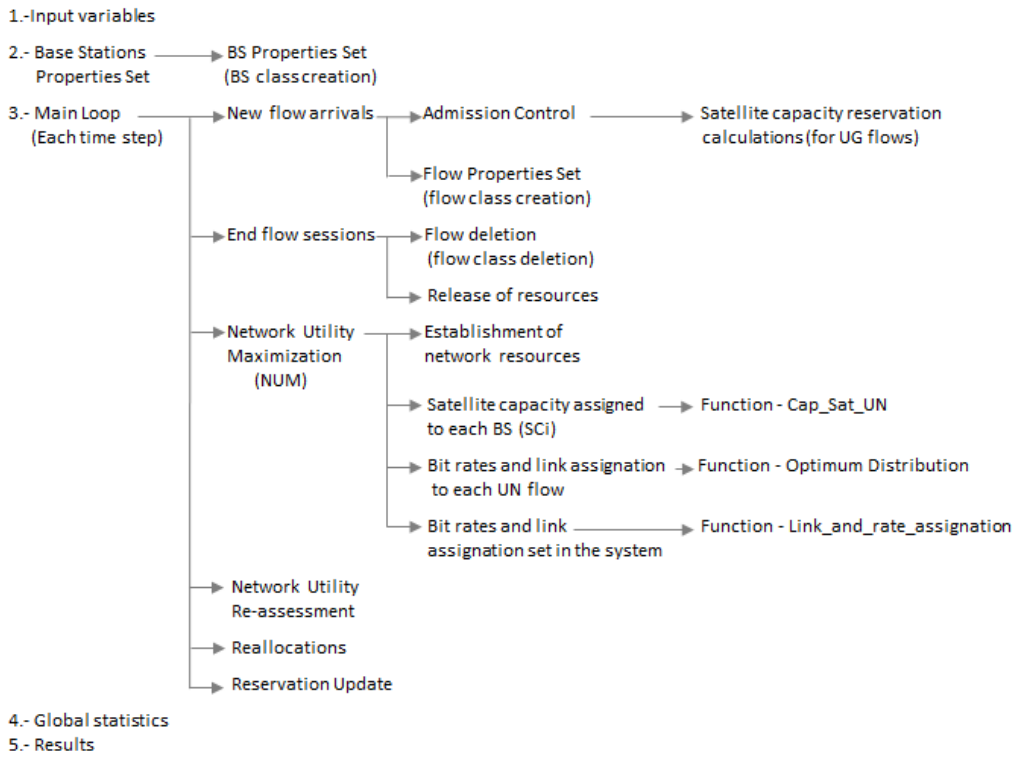

FIGURE B.4: General TE Application Structure.