INTERNATIONAL JOURNAL OF SATELLITE COMMUNICATIONS AND NETWORKING

Int. J. Satell. Commun. Network. 2016; **00**:1–46

Software Defined Satellite Cloud RAN

Toufik Ahmed[1], Emmanuel Dubois[2], Jean-Baptiste Dupé[2], Ramon Ferrùs[3], Patrick Gélard[2] and Nicolas Kuhn[2]

[1] *CNRS-LaBRI, University of Bordeaux, Bordeaux INP. tad@labri.fr*

[2] *Centre National d'Etudes Spatiales (CNES). firstname.lastname@cnes.fr*

[3] *Universitat Politècnica de Catalunya. ferrus@tsc.upc.edu*

SUMMARY

This paper provides a feasibility study on the virtualization of a DVB-S2/DVB-RCS2 satellite ground infrastructure and its SDN-based management and control. The proposed framework, SatCloudRAN, is expected to increase the opportunities of smoothly integrating the satellite components in forthcoming 5G systems. We analyze the design of SatCloudRAN by considering various chaining of virtual and physical functions and the characteristics of the links between them. We based our analysis on a generic architecture of bidirectional access networks that follows the normative documents of the broadband forum and leverage virtualization and softwarization technologies, namely NFV and SDN, to achieve a flexible and programmable control and management of satellite infrastructure. Using a SatCloudRAN approach, network operators will be able to provide: (1) optimized dynamic QoS, (2) resilient management of multiple satellite gateways, and (3) dynamic bandwidth on demand. Copyright © 2016 John Wiley & Sons, Ltd.

Received . . .

KEY WORDS: SDN; NFV; satellite networks; virtualization

1. INTRODUCTION

5G is not only about increasing the throughput or reducing the latency. The objectives behind this initiative are much wider and aim at providing Internet service anywhere, anytime and with any device [1]. To achieve this goal, the various access technologies shall be inter-operable, but each component should fulfill its own role, to provide ubiquitous seamless coverage. The role of the satellite in 5G is deeply discussed in [2]: the job of the satellite is mainly driven by its inherent strength, that are large coverage complementing the terrestrial coverage, the resilience that is necessary in critical telecom missions or the high broadcast throughput. This vision on the role of the satellite in the forthcoming 5G is yet to be revisited when deployment of the 5G is driven. It is necessary to assess the feasibility of integrating the satellite in future infrastructures to benefit from their natural advantages.

Satellite bidirectional access networks are of interests for many markets to: (1) provide reasonable Internet access in rural areas, where commercially viable broadband service may hardly be realized, (2) provide services anywhere and anytime including coverage to wide areas , and (3) broadcast data to millions of users. The interest in more cooperative interactions between satellite and terrestrial networks is not new [3–6], and some access providers start offering broadband bundles that conjointly use satellite and terrestrial resources such as the National Broadband Network (NBN) initiative.[*] Despite these initiatives, satellite networks might be difficult to assess for a terrestrial operator. There may not be common interfaces for resource management and control of terrestrial and satellite networks, since there is no convergence in their management planes. Moreover, the satellite ground segments exploit rather specialized functions such as tuned Transmission Control Protocol (TCP) proxies for satellite networks [7] or specific low layer mechanisms for the Digital Video Broadcasting - Satellite - Second Generation (DVB-S2) / Digital Video Broadcasting - Return Channel Satellite - Second Generation (DVB-RCS2) [8], which specificities are not known by terrestrial operators.

Anticipating flexible and standard control of satellite network resources will not only help towards a seamless convergence between satellite and terrestrial segments, but may also result in an increased service innovation and business agility. The recent years have witnessed a major shift towards Software Defined Networking (SDN) [9] and Network Function Virtualization (NFV) [10].

---

[*] See: http://www.nbnco.com.au/

These technologies are only identified as necessary key technical components of the 5G [1], so that 5G's requirements on flexibility and performance can be fulfilled. The introduction of SDN and NFV technologies within the satellite ground infrastructure along with the terrestrial network could pave the way for fully unified control plane that would allow operators to efficiently manage and optimize the operations of their terrestrial and satellite networks. We have analyzed, in our previous work [11], the opportunities and the challenges of using SDN and NFV in satellite networks. In particular, we have presented three scenarios that could provide some improvement areas through the introduction of SDN and NFV in the satellite ground infrastructure. In [11], we described our main target scenario, that is to let multiple tenants share an opened satellite ground segment infrastructure. This approach can be seen as offering wholesale access to satellite network resources along with customizable control and management of equipment as well.

In general, network virtualization involves the implementation of network functions in software that can run on a range of industry standard hardware [12]. Ubiquitous, convenient and on-demand access to a shared pool of configurable computing resources (*e.g.*, networks, servers, storage and services) can be rapidly provisioned and released with minimal management effort or service provider interaction. Virtualizing some functions that currently take place within the satellite gateways would improve the flexibility and the reconfigurability in the delivery of satellite network services. In the light of the increasing adoption of SDN and NFV technologies within terrestrial networks and the promised flexibility that would induce more interest in using satellite networks, this paper fills a gap in analyzing how to realize both (1) the virtualization of the satellite gateway and the satellite core network functions and (2) the management and control of a virtualized satellite ground infrastructure.

Based on a thorough analysis of the DVB-S2 and the DVB-RCS2 normative documents [8, 13–16], we provide a detailed analysis of (1) how the control functions are currently implemented, (2) how they can be virtualized and (3) how their management can be enhanced. Even if this analysis does not include a quantitative discussion of the advantages of SDN-control and NFV virtualization, we believe that it provides a valuable basis for a qualitative discussion. Indeed, it can be used to identify the aspects that have to be carefully considered in the virtualization process of a satellite gateway.

The main contribution of this paper is a novel framework named Satellite Cloud Radio Access Network (SatCloudRAN) that leverages cloud-based infrastructure and SDN-enabled network virtualization to deliver cost efficient, high-level resources availability and flexible resources sharing.

The rest of this paper is organized as follows. Section 2 describes (1) how the satellite network can be interconnected with the terrestrial network, (2) important functions that take place in the satellite core network and (3) the functions that are integrated in a satellite gateway. We propose in Section 3 a methodology to assess the feasibility of virtualizing the processes within the satellite network. In Section 4, we determine how the functions of a satellite gateway can be decomposed in a set of functions that could run as virtualized network functions and another set that would remain embedded in legacy hardware appliances. In the light of what processes can be isolated from each other, we assess the feasibility of virtualizing them in Section 5. Section 6 discusses the SDN control of a satellite core network with examples of controlled functions, such as bandwidth on demand or dynamic Quality Of Service (QoS). the impact of having a SDN controller that may be far away from the network element under its control. We conclude this paper in Section 7.

2. GEO BROADBAND SYSTEM

Currently, satellite Internet access are mainly done through GEostationary Orbit (GEO) broadband satellites, which ground segments systems are mainly proprietary. They follow nonetheless the spirit of the normative documents described in the DVB-S2 and the DVB-RCS2. Our analysis will be based on published DVB-S2 and DVB-RCS2 documents [8,13–16]. It is worth pointing out that our analysis can be extrapolated to other systems, such as Low Earth Orbit (LEO) constellations.

After a brief description of the main components found in a GEO broadband system, this section provides a description of the data, control and management plane functions that form part of a typical satellite gateway. The Table I presents the requirements for a GEO broadband system and the functions that are detailed in the rest of this section. The rationale is to clearly describe the key processes which virtualization will be further discussed in this paper. We base this description on our understanding of the DVB-S2 and DVB-RCS2 normative documents, such as [8, 13–16]. Even if this analysis focuses on the DVB system, it can be applicable to other systems.

Table I. Requirements for a GEO broadband system.

Requirement	**Rationale**	**Featured functions to fulfill the requirement**
Optimized spectrum efficiency	Expensive satellite resource	Fade Mitigation Techniques (FMT) Adaptive Coding Modulation (ACM)
		Physical layer recovery mechanisms
Connect terminals	Strong requirement for bi-directional access	Log-on procedure
		Service Level Agreement (SLA)
Share the capacity	Multiple terminals access	Synchronization
		Access gateway QoS
		Channel access methods
Enable and improve an end-to-end (E2E) connectivity	Network connectivity between components	Aggregation between multiple gateways (QoS)
		Connectivity to Broadband Network Gateway (BNG) (Ethernet, ipv4/6, etc.)
		Performance Enhancing Proxy (PEP) Security, etc.
Manage the network	Network management [15]	Fault management
		Configuration management
		Accounting management
		Performance management
		Security management

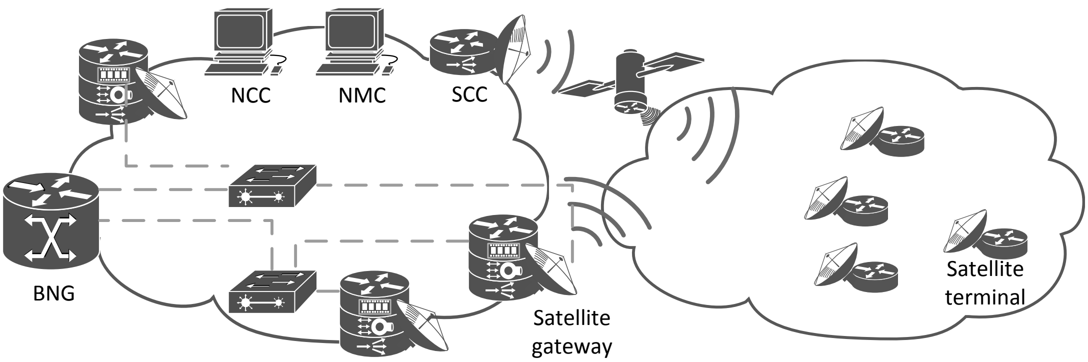

Figure 1. Satellite network architecture - source [17]

2.1. Satellite core network

In the context of satellite broadband access for fixed communications, a theoretical general reference model for a multi-gateway satellite ground segment is structured in several main subsystems, as depicted in Figure 1.

The "satellite access network" includes the satellite gateways and the satellite terminals, which are interconnected through the resource of one or several satellite channels. It can use a variety of network topology (star, multi-star, mesh or hybrid star/mesh) and provide a variety of types of connectivity.

The "satellite core network" is an aggregation network that interconnects different satellite gateways and includes the network nodes located at international Point of Presence (PoP)s to interconnect with other operators, corporations and Internet Service Provider (ISP). Typically, the satellite core network is built around an optical backbone with switching and routing equipment nodes based on Internet Protocol (IP)/MultiProtocol Label Switching (MPLS) or carrier grade Ethernet technologies. BNG can also form part the satellite core network if the satellite operator is a Network Service Provider (NSP).

The "control and management subsystems" is composed of Network Control Centre (NCC) and Network Management Center (NMC). NCC is used for real-time control of the connections and associated resources allocated to terminals that constitute one satellite network. NMC is used for non-real-time management functions related to a single satellite network. In addition, there is a

Satellite Control Center (SCC) in order to manage the satellite in-orbit platform and the satellite payload.

The reference architecture of a satellite gateway is depicted on Figure 2, which shows the following main elements that compose a typical satellite gateway: (1) an OutDoor Unit (ODU)[†], composed of an antenna and its radio components (Block Up Converter (BUC) to transmit to the satellite and Low Noise Block-converter (LNB) to receive from the satellite), we define here the satellite hub as the place where the ODU is located; (2) a physical gateway, dealing with physical layer related processes; (3) an access gateway, dealing with Media Access Control (MAC) layer related processes; (4) a network connectivity block, dealing with the interface for aggregation network access (IP router, Ethernet switch).

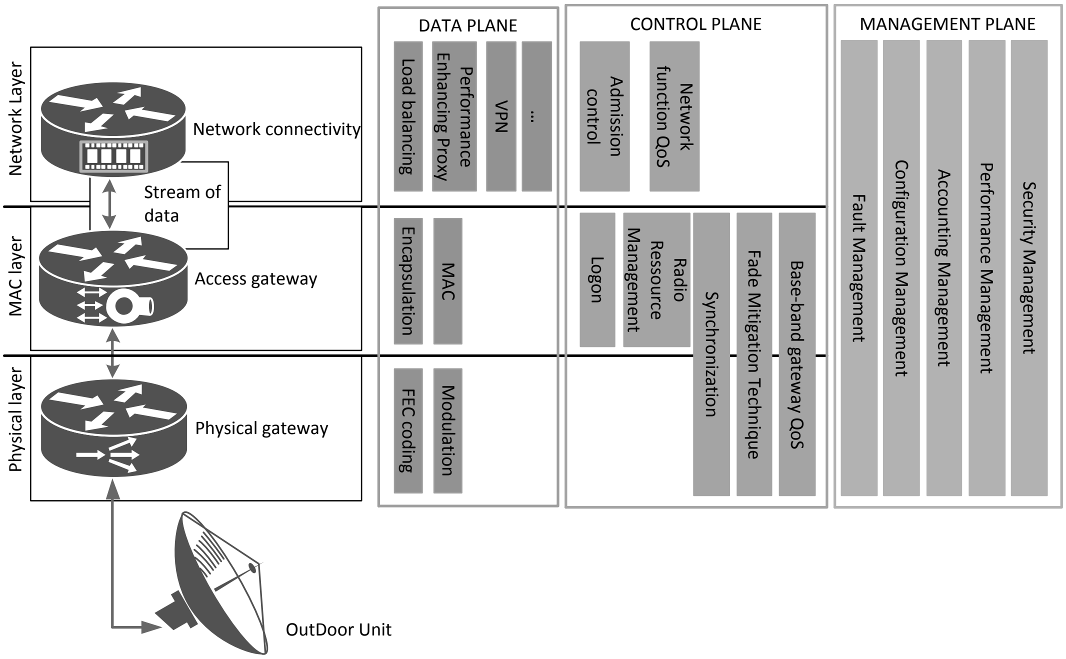

Figure 2. Satellite gateway reference architecture

There are various main processes within the baseband gateway. Each process is composed of set of functionalities that can be later isolated and virtualized when applicable. To support a discussion

---

[†] The ODU commonly refers to the satellite terminal; however, in general, it describes the equipment that is located outside of the building.

on their potential virtualization, we review in the rest of this section the main functionalities within each process.

2.2. Data plane processes

The data plane encompasses the actual transmission of IP packets on the satellite access network. Functions of the network connectivity, access gateway and physical gateway are presented hereafter.

2.2.1. Network connectivity In Figure 2, we drive a non-exhaustive list of processes, which could be considered as network functions. Other processes, such as data compression, firewall, or deep packet inspection, could have been considered but have been voluntarily omitted for the sake of clarity. This section quickly presents functions that could be part of the network connectivity of a satellite gateway: Virtual Private Network (VPN), load balancing and PEP processes.

A VPN [18] securely connects isolated computers or regional networks to each other and the head-office. The connection may be over the Internet, avoiding the setup of a private network: the traffic is encrypted and isolated via an IP Security protocol (IPSec) tunnel between the origin and destination network. In this point to point solution, VPN users are authenticated to securely access remote resources such as email and Intranet just like if they were on the central network. No private information is visible and message integrity is guaranteed. The VPN is a tunneling mechanism based on IPSec, which offers authentication and encryption.

The load-balancing process can be seen as a process that deals with spliting the load over various paths to simultaneously exploit the capacity of different links. As example, load-balancing techniques can be used to share the load between various carriers, various gateways or various access technologies (terrestrial/satellite). This function is an example Distributing traffic arbitrarily among the available links despite their different characteristics or their current load can result in sub-optimal performance. This paper focuses on one example of load-balacing: a multipath routing entity such as the Hybrid Customer Premises Equipment (HCPE), which can seamlessly select the appropriate access technology for the different types of network traffic, requires the ability to know the QoS requirements (*e.g.* bandwidth, latency and packet loss) of each particular traffic flow and to compare them with the real-time status of each of the available access links. Dealing with the heterogeneity of the available links is an open issue in multipath activities.

The PEP provides a combination of compression, caching techniques and TCP acceleration. Due to TCP performance degradation over satellite links, PEPs are currently the most commonly adopted solution to achieve good transport performance (in terms of link utilization and user experience) whatever the available TCP stack at both ends (clients and servers). The location of PEP terminations in the architecture has important impacts on the overall network design. TCP session interception is necessary for acceleration and data compression, with each session transparently split into three TCP sessions. Splitting the session in three segments enables local acknowledgement of session establishment and data, without the impact of round-trip delays. It also enables faster ramp-up of TCP throughput on LAN segments and faster recovery to packet losses occurring on the Local Area Network (LAN). However, this TCP proxy mechanism remains fully transparent for clients and servers, as well as the devices on the path of the middle TCP session.

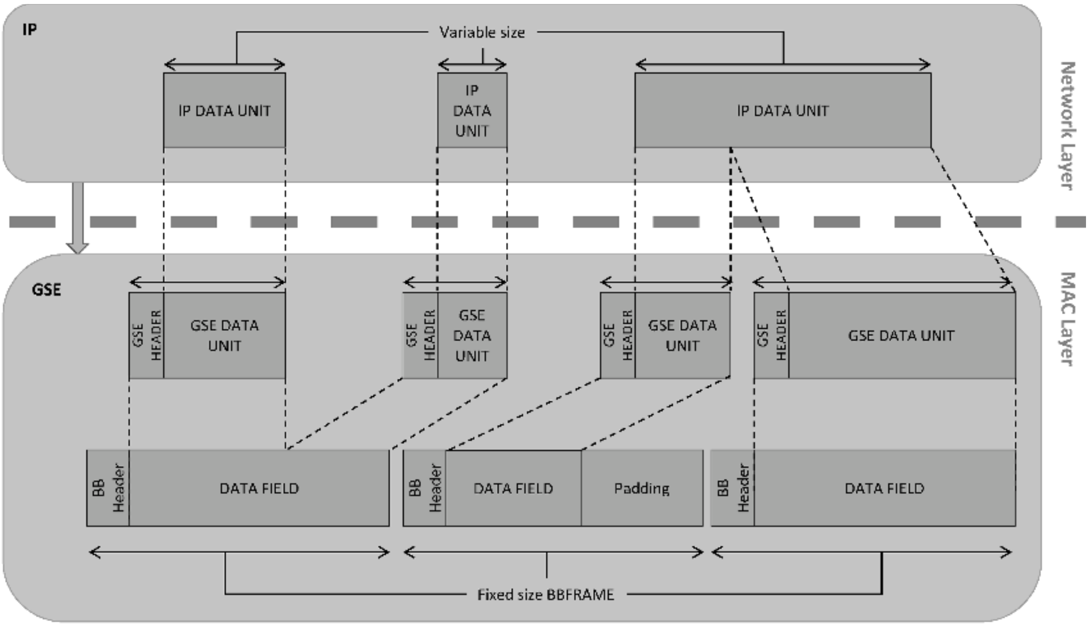

Figure 3. MAC layer data plane architecture on the forward link

2.2.2. Access gateway The protocol stack of the MAC gateway's data plane is given on Figure 3, on the forward link. MAC layer encompasses various functionalities, such as encapsulation and fragmentation, medium access control itself, protocol multiplexing, scheduling, QoS, addressing scheme and errors detection. The present section focuses on the description of the data plane, whereas processes related to the control plane are detailed in Section 2.3.

Encapsulation on DVB-S2 forward link is performed by Generic Stream Encapsulation (GSE), which is responsible for encapsulating and adapting the upper layer's data to transmission on physical layer frames, known as BaseBand FRAME (BBFRAME)s. As illustrated on Figure 3, GSE deals with various sizes incoming packets and must generate BBFRAME which size depends on the current physical layer coding. On the return link, encapsulation is performed by Return Link Encapsulation (RLE) which features with a similar process. It is worth pointing out that RLE provides more functionnalities, since it can operate at different levels, such as frames level, Protocol Data Unit (PDU) level. The paradigm of both RLE and GSE encapsulation techniques is slightly different: while GSE multiplexes packets in larger frames, RLE splits incoming packets into smaller frames.

Concerning medium access control, the processes are very different on the forward and return link. On the forward link, the access method used is Time-Division Multiplexing (TDM), where all user data is transmitted sequentially on a single carrier. This mechanism is simple in the sense that there is no resource sharing to handle on the physical level: multiplexing of user data is done at MAC level.

On the return link, the satellite resource is shared among the users. Two access methods can be distinguished: on-demand access or contention access. In the former, a terminal receives dedicated resources on its own to communicate with the gateway. In the latter, some resources are reserved for contention access, where several terminals can compete to obtain the resource. Dedicated access, which is more common in currently deployed systems, can be through a Demand Assigned Multiple Access (DAMA) mechanism, while contention access techniques are usually based on Slotted Aloha (SA) and its numerous derivatives.

2.2.3. Physical gateway The physical gateway is responsible for the actual transmission of BBFRAMEs or reception of return link timeslots bursts bytes. The architecture of the physical gateway is depicted, for the forward link, on Figure 4.

Since the process is similar on the forward and return link, we focus here on the former. The physical gateway forwards “ready-to-sent” L-Band signal to the ODU with the following steps.

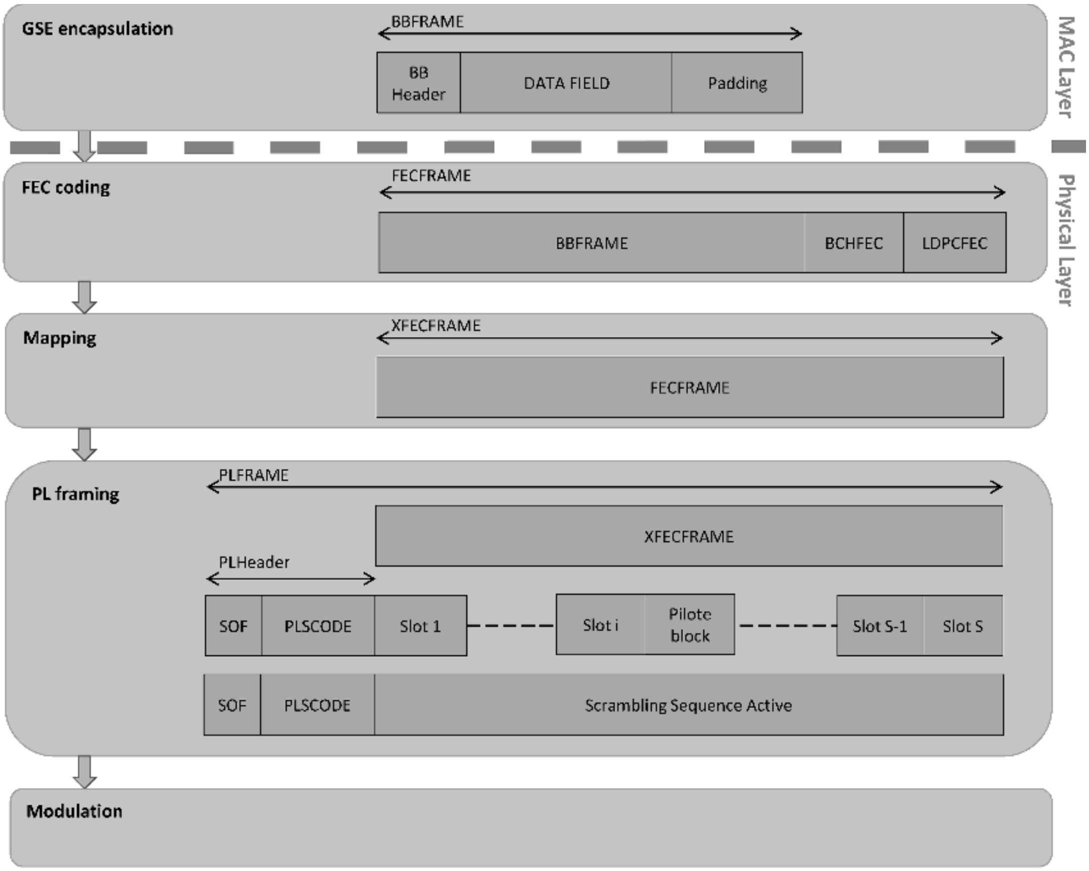

Figure 4. Forward link physical layer data plane architecture

Forward Error Correction (FEC) coding is introduced on the BBFRAME to detect and correct bit-errors that may occur during the transmission of data. Two FEC schemes are applied on the BBFRAME (Low Density Parity Check (LDPC) and BCH (BCH)) to improve the reliability of the transmission. The FEC coding block generates a FEC Frame (FECFRAME). FECFRAME have a constant size of 64800 or 16200 bits, yet because the amount of redundancy bits can evolve, the actual size of the BBFRAME is not fixed. As a result, the payload of a BBFRAME is defined by the FEC coding rate.

The mapping block maps the bits of the FECFRAME to complex symbols, and produces a sequence called the compleX FEC Frame (XFECFRAME). As the DVB-S2 and DVB-RCS2 standards define several modulations, the number of bits per symbol (modulation order) is variable, hence the transmission time of a given BBFRAME depends on the modulation. The association of coding rate and modulation order thus fully defines a BBFRAME and is called the MODulation and CODing (MODCOD).

The Physical Layer Framing (PLFraming) block generates a Physical Layer Frame (PLFRAME). The Physical Layer Header (PLHEADER) is composed of the Start-Of-Frame (SOF) and the Physical Layer Signalling CODE (PLSCODE). The SOF is used by receivers to identify the start of the PLFRAME. The PLSCODE indicates the MODCOD used, and the location of pilot symbols.

At the gateway level, on the return link, the physical gateway receives the signal from the ODU, synchronizes on the Multi-Frequency Time Division Multiple Access (MF-TDMA) frame, demodulates depending on the modulation, decodes and forwards the frames to the access gateway. This simplified view is related to the data plane only: measurements are done, so that information on clock drifts or Signal-to-Noise Ratio (SNR) on the forward link can be forwarded to the control plane, so that processes can be adapted. More information on that aspect can be found in Figure 13. These exchanges on data and control planes imply interfaces between access and physical gateways.

2.3. Control plane processes

Control plane process includes all the processes that set up the necessary procedure for data to be forwarded across the satellite network. In the routing area, the control plane is responsible for choosing the optimal routes and indicate routers how to actually forward packets from one point to one other.

These processes are mostly in the access gateway but can take decisions that will be applied at the physical gateway: the control information can either be carried out along with data packets, or through specific interfaces. Control processes mainly deal with deciding the physical gateway parameters that should be used, such as choosing the FEC coding rate, the modulation to be used, the moment at which synchronization messages shall be transmitted. Thus, they parameterize the processes in the data plane that are shown in Section 2.2.

2.3.1. Logon In order to access the radio resources and request capacity, a terminal has to gather the necessary information to communicate with the gateway. This can be summarized in two phases : forward link procedure and logon.

The forward procedure is for the terminal to receive signaling information from the gateway. The information that the terminal can obtain during the listening process is related and not limited to: (1) the satellite and its gateway; (2) the superframe sequence number; (3) the satellite ephemeris; (4) the logon timeslots.

The terminal can exploit the logon timeslots to logon. This phases is related to the synchronization process, since terminals require to be synchronized to accurately locate this slots. Indeed, the transmission of actual data burst cannot start until both terminal and gateway are synchronized.

2.3.2. Synchronization For the terminals to accurately share the satellite resource, each of them need to be synchronized with the gateway. The accuracy of frequency and time synchronizations is important to guarantee, as one example, that terminals respect their given slots to transmit data on the return link.

Terminals and the gateway use internal clocks that independently drift over the time. On top of these clock drifts, the global clock synchronization is challenged by the fact that the satellite is moving and terminals are not located at the same place, which results in different jitters and different Round-Trip Time (RTT)s for each terminal.

On the frequency synchronization: DVB-S2 Forward Link frequency synchronization is dealt by using the pace at which the SOFs are received. At the terminal, the Directed Digital Phase Locked Loops (DD-PLL) loops with phase error detection symbols sequence on the received symbols to improve frequency synchronization.

On the temporal synchronization: To achieve a global clock synchronization between the different terminals and the gateway, the gateway frequently transmits a timestamp Network Clock Reference (NCR) that is exploited by all the terminals on the return link. The timestamp records the transmission of the first symbol of the PLFRAME N and insert it in the $N + 2$ BBFRAME, where a slot had been dedicated for it.

When terminals are able to locate the logon slots accurately, they can send logon requests. Knowing the slot in which the logon burst was sent and its reception time, the gateway can transmit correction messages to the terminal. At the receiver level, the transmission of data burst cannot start until the correction messages are not “close to zero”.

When the terminal is logged on and synchronized, the synchronization is regularly monitored. When no NCR has been received during a certain period (which is implementation dependent) or when there is a loss in the synchronization (which can be measured at the gateway), the terminal shall cease transmission.

2.3.3. Radio Resource Management Radio Resource Management (RRM) encompasses techniques needed to distribute the available frequency bandwidth in order to allow bidirectional communication between the terminals and the gateway. As shown in Figure 5, we illustrate the example of three Satellite Virtual Network Operator (SVNO)s sharing the frequency resource which is divided between the forward and the return links. The relevance of this example depends further on the role model.

On the forward link, the gateway uses all the available bandwidth, possibly with several carriers, to communicate with the remote terminals. A carrier is a single TDM where packets addressed to terminals are multiplexed within the BBFRAME as shown in Figure 3. This scheme requires little signaling, apart from the allocation of a terminal to a carrier. The carrier settings (frequency bandwidth, symbol rate) are usually set once, and terminals are assigned to a carrier that suits their bandwidth needs. Hence, in terms of RRM, the forward link is mostly static for the assignment

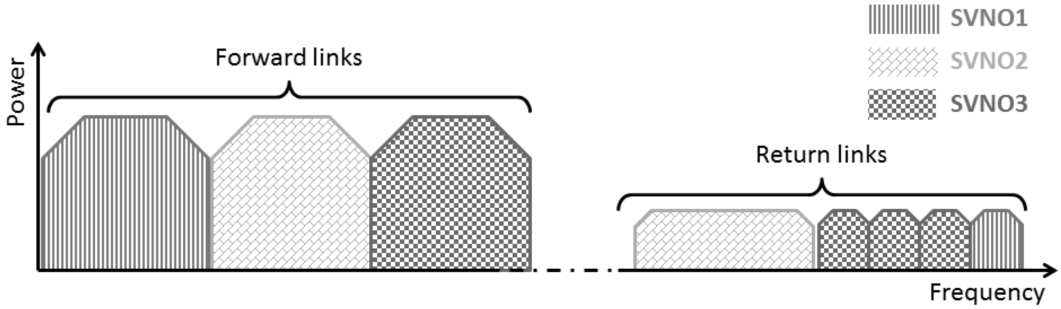

Figure 5. Example of the sharing of the spectrum between SVNOs. The way the frequency is shared among the SVNOs is more related to the management process; the scope of this figure is to show how the frequency can be shared between the forward and the return links

of terminals to carriers. The scheduling of data within the forward link is mostly dynamic and we consider in this article that the scheduling is dealt with the baseband gateway QoS.

On the return link, the goal of the RRM is to distribute the resource between the terminals to let them communicate with the satellite gateway. The access method proposed within the DVB-RCS2 standard consists in dividing the available bandwidth in small time-frequency units called Bandwidth-Time Unit (BTU). Contiguous BTUs can be grouped into a timeslot, several timeslots form a frame, and several frames are themselves grouped into a superframe. This hierarchy is described on Figure 6.

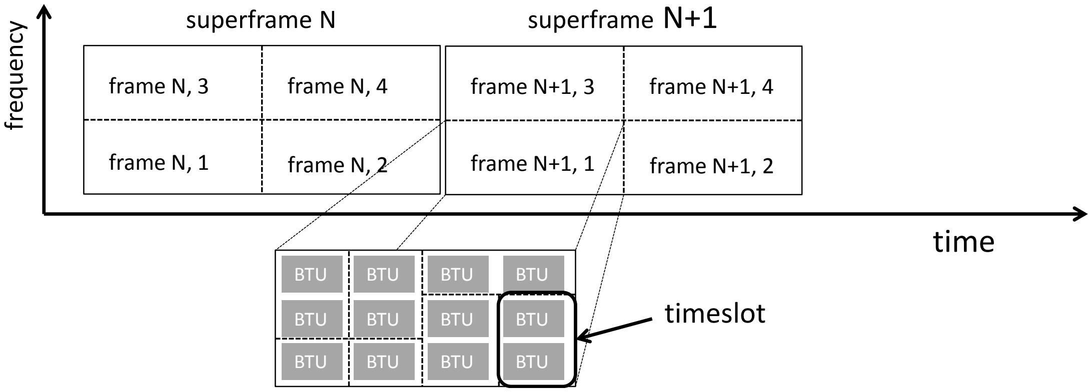

Figure 6. DVB-RCS2 return link hierarchy

This hierarchy allows an efficient resource allocation: each terminal can be assigned by the gateway a timeslot of its own, in dedicated access, in order to transmit data without competing with other users. A timeslot is defined by the number of BTU it is composed of, as well as the modulation, coding rate and payload the terminal has to use. A timeslot can also be dedicated to contention access, and can thus be used by any terminal wishing to transmit data without having to explicitly ask the gateway for it.

Terminals can send periodically to the gateway a traffic request, expressed either in rate (Rate-Based Dynamic Capacity (RBDC)) or in volume (Volume-Based Dynamic Capacity (VBDC)). The request may be only sent in the SYNC slot, but vendors can also use in-band request. If this update process is done at each superframe, only a subset of all terminals can actually update their requests, in order to minimize the overhead created by the control plan on the return link. At each superframe, the gateway collects all the terminals requests and allocates the next available superframe to terminals, following their requests. This process is shown, for a single terminal, on Figure 7.

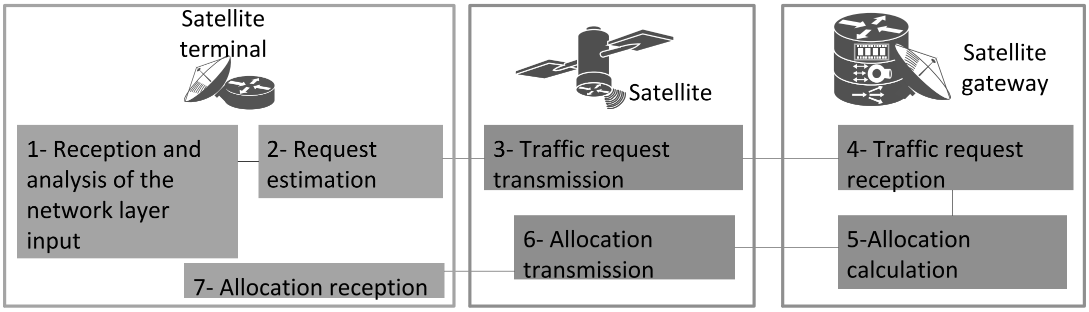

Figure 7. DVB-RCS2 Allocation process

Two algorithms drive the performance of resource allocation: request estimation and allocation calculation. Request estimation is done each time a terminal needs to update its request. In most systems, this period can be typically over one second, which is very long considering the dynamics of traffic on return link. Hence, the process of estimating traffic needs is a very difficult task: it has to match the needed resource as much as possible. On the other side, allocation calculation has to carefully trade between spectral efficiency, QoS and respect of users' needs.

A noticeable feature of this system is the duration of the allocation cycle, totalizing at least one second. This duration is partly caused by the very large propagation delay inherent to the altitude of the satellite in geostationary systems. Making an efficient use of the available bandwidth becomes particularly challenging in this context, and influences deeply the complexity of resource allocation and request estimation.

2.3.4. Fade Mitigation Technique First generation of Digital Video Broadcasting - Satellite (DVB-S) systems were designed for a worst-case scenario, where attenuation was considered maximum. The robustness of the transmission towards errors, controlled by both coding rate and modulation order, was adapted to provide a Quasi Error Free (QEF) link even in the worst case. It could therefore provide a very high availability (up to $99.6\,\%$ of time) but was largely oversized for most of the time. To better utilize the medium, DVB-S2 and DVB-RCS2 systems introduce adaptive modulation and coding schemes that consider the current channel quality.

On the DVB-S2, the key concept of this technique is to monitor link quality in real-time, with the help of known symbols sequences, included along regular packets, on which an estimation of the current SNR can be done. Then, this estimation is send back to the transmitter who can adapt its coding rate and modulation order to best fit the actual transmission conditions.

This process is shown on Figure 8 with focus on the forward link, where it is called ACM. The ACM process usually sets a target Packet Error Rate (PER) as a reference, such as 10^{-7} for the forward link. The MODCOD is chosen at the access gateway level but is not carried in the BBFRAME header but as a single message until it can be included in the PLSCODE at the physical gateway level. In terms of signaling, the only information actually transmitted is the SNR estimation reported from the terminal to the gateway. The overall process is controlled at the MAC level, within the gateway, and its periodicity is the same as request sending, typically one second or more.

On the DVB-RCS2 link, DVB-RCS2 can also feature return link FMT, with the help of known symbols, or pilots, included in bursts. Once the gateway receives a burst, it can estimate link quality with those pilots, and adjust the MODCOD used in the timeslots allocated to the terminal. The gateway may also adapt the time-frequency distribution, considering the channel conditions. This process does not involve additional explicit signaling, unlike forward link ACM, and is not

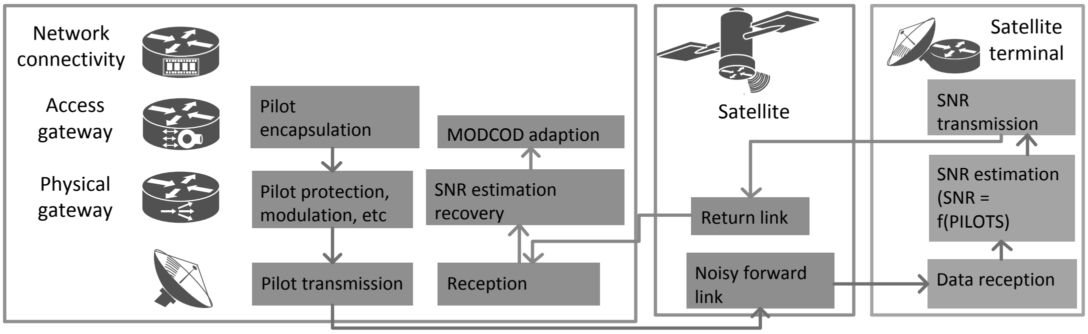

Figure 8. FMT mechanism on forward link

mandatory. Considering the amount of possibilities, the allocation could be mainly from static set of carriers with constant MODCOD to highly variable carriers with time slots with different MODCOD; the decision on which possiblity to consider for a given system is out of the scope of the FMT process, but needs to be consider in the design of the system.

2.3.5. Terminal Admission Control The terminal admission control is a method that can be used to restrict the access to a network. If a network device has been configured to consider admission control, it may force user authentication before granting access to the network.

An application running on the terminal connects to the network by a logical session. Based on the protocol used, subscriber sessions are classified into types that depend on whether the interconnectivity is dealt with at layer 2 or layer 3.

During the Authentication, Authorization, and Accounting (AAA) process, (1) subscribers are authenticated before establishing a subscriber session; (2) authorize subscriber are authoerized to access specific network services or resources; (3) usage of broadband services is tracked for accounting or billing.

2.3.6. Control plane QoS The QoS is the capability of a given network to carry out a given data flow in good conditions (in terms of delay, jitter, loss rate, capacity, etc.). Thus, as shown in Figure 2, the QoS optimization process operates at both the access gateway and the network functions levels. This process can be divided into two sub-processes: the “network function QoS” and the “baseband gateway QoS”.

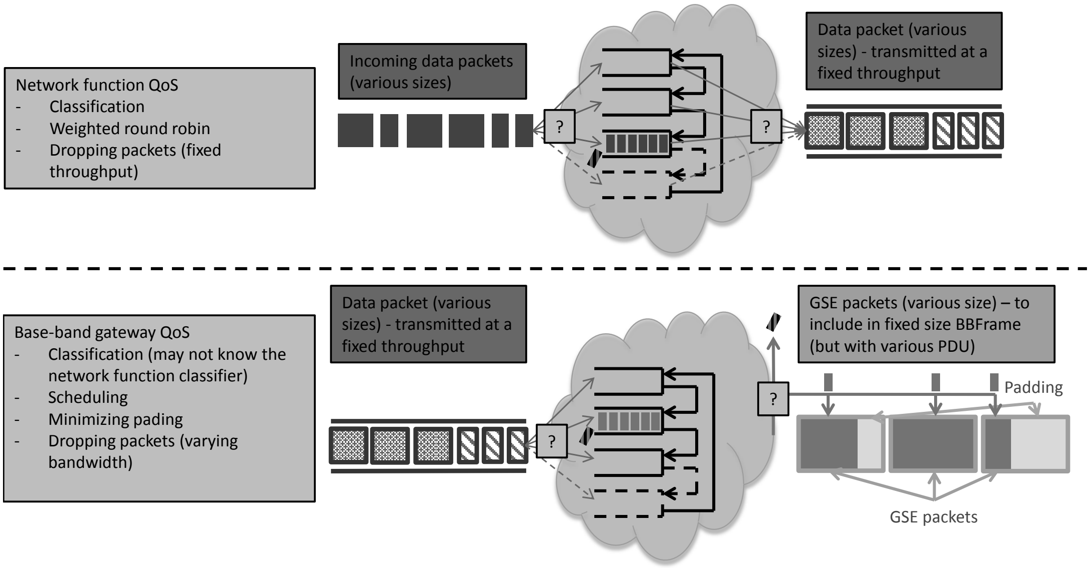

Figure 9. Overview of the different operations done at the network function QoS and the baseband gateway QoS

The combination of the network function QoS and the baseband gateway QoS deals with variously sized incoming packets, such as shown in Figure 3, and a limited resource as shown in Figure 5. Figure 9 sums up how the network function QoS and the baseband gateway QoS cab interact.

The network function QoS classifies the packets in dedicated sub-queues depending on defined parameters (tag, flow, packet size, etc. . .): the objective is to adapt scheduling algorithm so that packets are dequeued with considerations of the requirements for each given class of traffic. Network-level QoS is usually responsible for multiplexing the incoming flows before passing them to the access layer. When a sub-queue is full, incoming packets are dropped. The classifier may consider the nature of the transport protocols in its classification, so that the sub-queues containing flows that are reactive to congestion can consider Active Queue Management (AQM) [19] techniques to reduce the buffering and the latency. Since they cannot access to the lower layer characteristics, some implementations of network function QoS consider that the whole available goodput can be exploited by the lower layers: this case is referred as the “clear sky” case. Because the throughput is sensible to channel quality (due to FMT), the lack of information can lead to overflows. To sum up with, the network function QoS: (1) considers classes of traffic, each of them having specific requirements; (2) classifies the incoming traffic; (3) applies a scheduling algorithm for selecting the packet to dequeue; (4) considers a “clear sky”; (5) may drop packets. Several QoS architectures have been proposed, such as Diffentiated services (Diffserv) IP QoS [20] or Metropolitan (Metro) Ethernet QoS [21].

As shown in Figure 9, data packets of various sizes enter the baseband gateway with a fixed throughput. The baseband gateway generates variously sized GSE packets that are to be included in BBFRAME. The payload in a BBFRAME depends on the MODCOD. Joint algorithms on the decisions of the scheduling and the available payload have been proposed [22]. The baseband gateway QoS must therefore deal with fixed throughput incoming data and variable available throughput for outgoing packets, while minimizing the amount of padding and the number of dropped packets. The incoming traffic can be classified and re-organized in a variable number of sub-queues, which usually corresponds to a mapping between QoS defined at the network level and access level. The purpose of the scheduler on the baseband gateway is to ensure that it can cope with throughput variations without affecting traffic QoS. The decision operates with inputs from the FMT and from the RRM.

In Figure 2, the baseband gateway QoS is shown to be at both the MAC and the physical layers, since information from the physical layer of the return link may be exploited by this process.

On the return link, the baseband gateway QoS is achieved in the process of the assignment of slots to the terminal and this decision is taken within the baseband gateway. The return link QoS includes the RRM, since it depends on the way to bandwidth is divided in multiple carriers for the return link.

2.4. Management plane processes

The non-exhaustive list of management plane processes that is shown in Figure 2 features the fault management (collecting data from various equipment to handle alarms or to detect and correct troubles), the configuration management (equipmenet configuration, device discovery, network provisioning), the accounting management (service billing), the performance management (collect error logs) and the the security management. More information of these processes can be found in [15, § 8.1.1], from where this list has been extracted.

This article has focused on the state-of-the art of control and data plane but management plane is also a major topic to consider in details in virtualization trend. The rationale of mentioning it in this article is to highlight this issue for potential future work.

3. ROADMAP TOWARDS THE DEPLOYMENT OF SATCLOUDRAN

In this section, we present the Cloud Radio Access Network (CloudRAN) approach, that is the trend in virtualizing the terrestrial mobile access. We also show how our proposed approach, the SatCloudRAN matches the virtualization process of the CloudRAN.

3.1. CloudRAN, the trend in virtualizing the terrestrial mobile access

If the entire burden of supporting high volumes is pushed to mobile network this would require operators to upgrade the capacity of their infrastructures by several orders of magnitude. These infrastructures have been traditionally based on a complex set of interconnected proprietary hardware appliances running different types of protocols and requiring specialized vendor-specific configuration tools. Furthermore, the cost for infrastructures in terms of deployment of mobile Radio Access Network (RAN), setup and operation is high enough to discourage any new hardware investment. It is therefore difficult to scale the network deployments for each situation, considering the cost and complexity constraints. The costs for backhaul from mobile base station to Evolved Packet Core (EPC) represent a significant part of operator revenue. As operators constantly introduce new sites and increase the number of base stations, the power consumption gets a dramatic rise [23]. Besides this, the introduction of new service would require a new specialized hardware and software to be installed.

To address the above-mentioned issues along with capacity, coverage, power consumption and upgrade, mobile operators are defining new architectures with centralized capabilities and service virtualization namely Centralized-RAN or CloudRAN.

This cloud-based centralized processing is a promising approach that aims to favor efficient operation, lower power consumption, provide agile traffic management, and improve network reliability. We acknowledge that these objectives may not all be granted, but future work could validate the fulfillness of these objectives. Further, it would enable to stimulate service innovation and reduce time-to-market to deploy new services. CloudRAN was a result of collaboration between Intel and China Mobile [23, 24] and is also of interest for other actors [25, 26]. This latter has conducted numerous trials and is expected to incorporate CloudRAN in its commercially deployed networks in China between 2015 and 2016 [27].

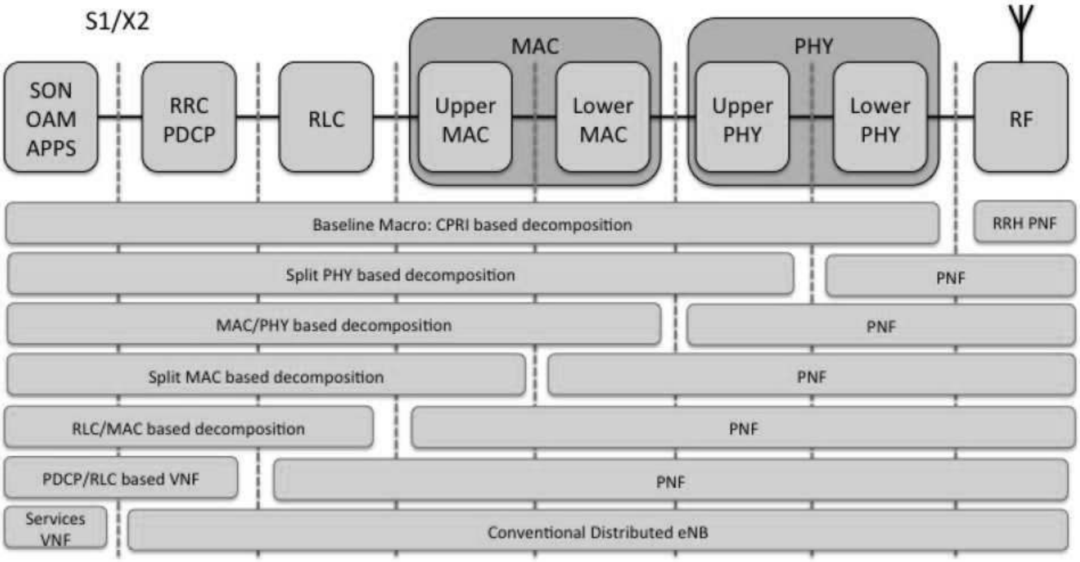

Figure 10. Possible ways to decompose the small cell - source [28]

Several scenarios have been proposed by the Small Cell Forum (SCF), where a certain segment of the Small Cell (SC) can be decomposed and virtualized as it is presented in Figure 10 [28]. As we move from left to right the decomposition becomes higher and the remote node that represents the physical entity becomes smaller. The rationale for examining these alternative splits is related to the associated requirements on the transport network for supporting the fronthaul link between the Virtual Network Function (VNF) and Physical Network Function (PNF) components. As an increasing set of functions are implemented as a virtual network function, the transport requirements in terms of bandwidth and latency become more onerous.

3.2. The SatCloudRAN

SatCloudRAN platform implements the separated baseband functionalities in a centralized cloud-based processing platform. This separation between the virtualized and the physical components can be achieved at various layers of the satellite architecture model such as the network layer, the MAC layer, the physical layer or up to the Radio Frequency (RF) front-end of out-door unit.

It seems worth pointing out that our proposed approach to virtualize the satellite network shows a high level of similarities with the approach that conducts the current virtualizing of terrestrial RAN networks. This point is however leveraged by the fact that network architectures in terrestrial and satellite systems are quite different.

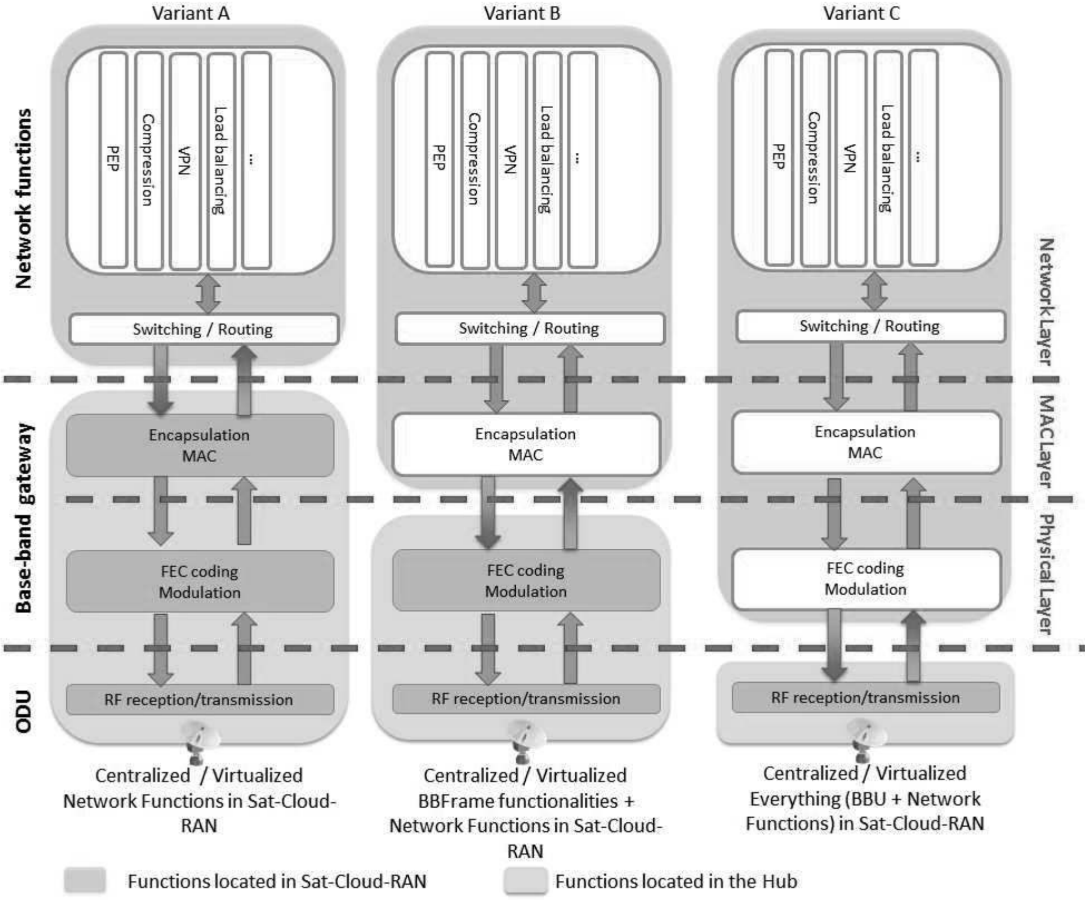

Figure 11. Variants for the functional split

In Figure 11, we show three different separation variants (A, B and C). The main difference between those variants is the distinction between the functions that would remain located in the satellite hub and those that would be moved to the centralized and/or virtualized infrastructure. Additional alternative decompositions where the split is made within the physical or within the baseband gateway functions could be relevant, depending on the conclusions of this study.

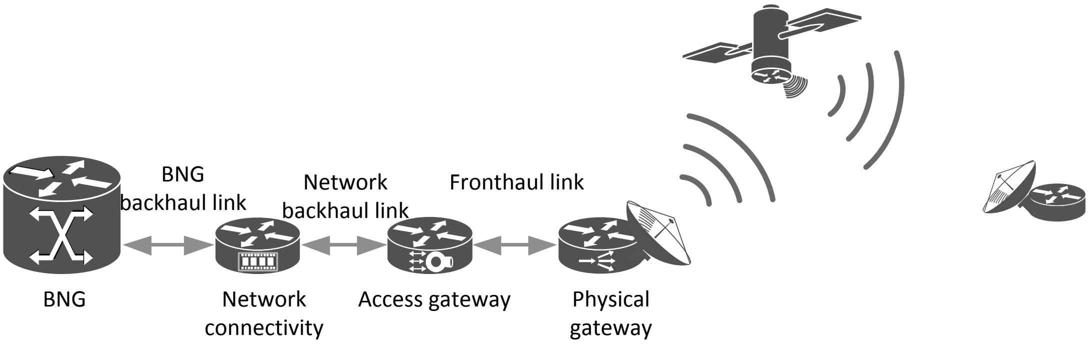

Figure 12. Fronthaul and backhaul

In Figure 12, we present the fronthaul design which is impacted by the process of separating functions. The fronthaul link is defined as the link between the physical gateway and the access gateway.

3.3. Roadmap towards the definition of the SatCloudRAN

Passing from a non-virtualized environment to the SatCloudRAN requires a specific roadmap: (1) identify the functions that can be isolated from the gateway and discuss their potential centralization (Section 4); (2) discuss the virtualization of the separated functions (Section 5); (3) assess the SDN control of the functions (Section 6).

3.4. How the SatCloudRAN can help in opening satellite system to new coming operators

Satellite Network Operator (SNO)s are looking for new business models to increase their customer base and extends the reach of their services offering. They are moving towards opening their infrastructure to be shared by multiple tenants, such as SVNO, and offering pay-per-use models instead of single-owned and used infrastructures. Multi-tenancy in infrastructure sharing model enables multiple tenants to cohabitate while being assured they can manage their own space in an isolated, flexible and secure fashion.

The SVNO model has emerged over the last few decades as many efforts have been made to open satellite system to new coming operator that can share cost and infrastructure with a host network operator. Different levels of granularity for controlling satellite system are already proposed [29]. The "managed services" offers a first step toward network control and bandwidth management for service provider who want to have a certain control on underlying resource provided by the satellite operator. The "SVNO model" allows a virtual network operator to get leased bandwidth with partial hub infrastructure control and management from the hosting satellite operator. SVNO can perform service provisioning, common network operation, and has full control of its own slice of network and end user. The "hub colocation model" (full SVNO) allows a SVNO to co-locate hub infrastructure in its teleport allowing greater control of the installed network equipment.

The aim of this separation is to enable the creation of an environment with fully virtualized capabilities allowing flexible management, installation, maintenance and operation of resources and services. This would thus facilitate the integration of satellite network in a hybrid network as a virtual layer infrastructure. Therefore, the proposed SatCloudRAN, presented in Section 3.2, helps SVNO providers get to market faster and at lower cost while gaining advanced control, more flexibility, and programmability of its allocated resources.

4. FUNCTIONAL SEPARATION OF PROCESSES

In this section, we present the functional separation of processes that take place within the satellite core network and the satellite gateway.

4.1. Fronthaul link characteristics and control plane processes

4.1.1. Variant A In the case of the variant A, the data packets to be forwarded are IP packets, and thus, there is no specific issue in carrying them out on the aggregation network and manage the connectivity between the gateways and the BNG.

The control processes that take place in this variant are the admission control and the network QoS. The network level QoS adds packets to a specific sub queue at a speed that is related to the network underneath and to the rate of incoming packets. Thus, the fronthauling link has not only a direct impact on the rhythm at which the packets arrive at the baseband gateway, but also on the relevance of the ordering of the incoming packets which depends on their class.

4.1.2. Variant B With the variant B, the data packets that are forwarded to the physical gateway are BBFRAME. In the case of no padding, directly carrying out fixed packets would only add the GSE header to the IP data packets. The connectivity between the access gateway and the physical gateway can be ensured by the use of layer 2 network segregation techniques.

The interactions between physical and MAC layers may however question the feasibility of this variant. Indeed, we present in Figure 13 the interaction between the control processes at these two layers.

For the logon process, the terminal applies a timer to retransmit its logon burst in case no acknowledgement has been received. The number of logon burst retransmission is limited. The norm does not detail the possible values for both the timer and the maximum number of trials, which are both implementation-dependent.

For the synchronization process, if the fronthaul link exhibits jitter, the NCR may not be transmitted at a fixed rate. Whatever the jitter in the fronthaul network, according to the normative documents, the NCR shall be updated at least ten times per second. On top of the issue related to the jitter in the fronthaul network, the potential losses of the BBFRAME carrying an empty slot for the NCR is a critical issue. If the terminal considers the NCR to be lost, it shall cease the transmission of data until it is synchronized again.

For the FMT process, if the precision of the measurement itself is not impacted by the splitting of variant B, the interactivity of the mechanism could be degraded by a fronthauling link introducing an important delay. This can result in selected MODCOD not matching the target PER, which would

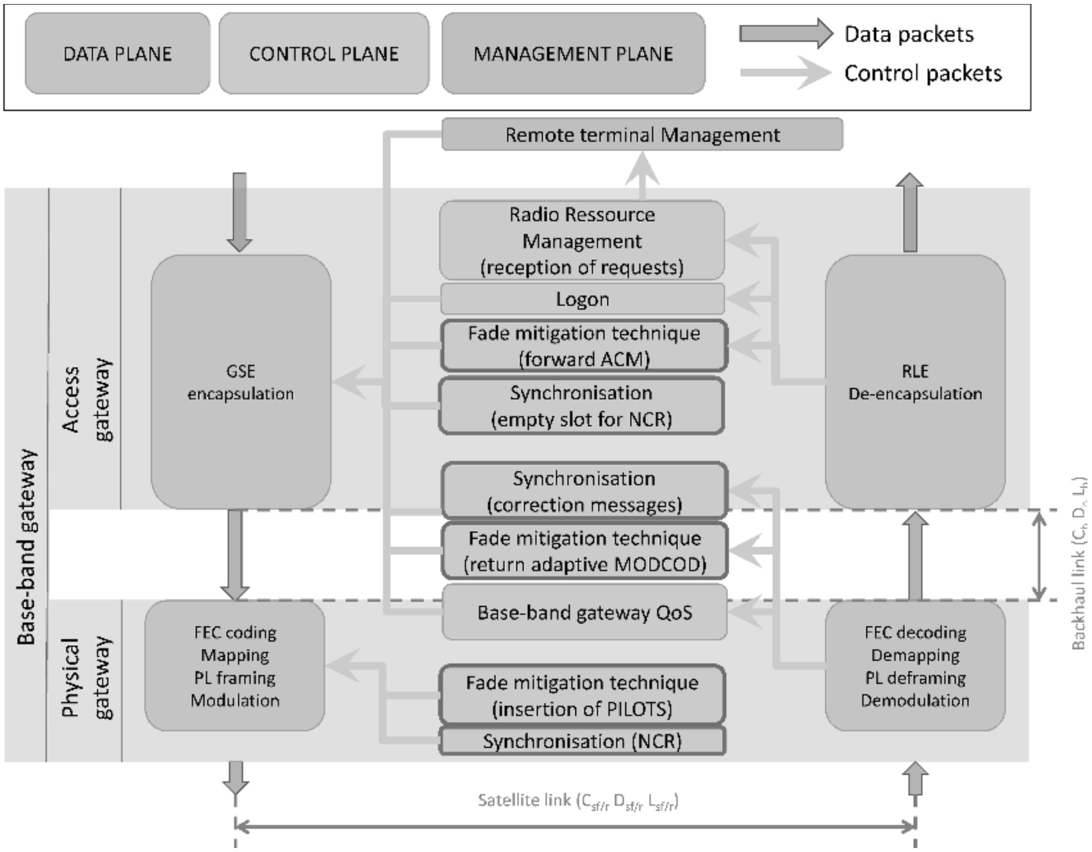

Figure 13. Interaction between the processes in the physical and the access gateways

lower overall performance. It is worth pointing out that there are ACM margins that may avoid this to happen.

If the fronthaul link between the physical and the access gateways shows a high PER, introduces a non-negligible amount of delay, or introduces jitter, there may be destructive impact on the logon procedure, and the synchronization process. It is worth pointing out that it is easier to target a low PER than overcoming latency issues. Moreover, this may also result in a non-adequacy between the reported SNR and the actual channel conditions and thus, misusing the expensive satellite resource in an efficient manner.

4.1.3. Variant C With the variant C, the data packets that are forwarded to the ODU are either PLFRAME (I/Q symbols) or directly the L band. The connectivity between the access gateway and the ODU can be ensured by the use of layer 2 network segregation techniques. In that context, the dedicated fronthaul channel could use the Common Public Radio Interface (CPRI) between the SatCloudRAN and the ODU.

The I/Q symbols may be transmitted over an optical network, which would ease the fulfillness of the requirements in terms one-way delay, jitter, throughput and bit error rate, as opposed to variant B. The L band signal could either be digitized or analogically transmitted.

If the L band is digitized, there is no specific need for a dedicated network. However, the choice of the fronthaul network would have an impact on the resulting feasibility of this solution, since the bandwidth requirements for digitizing the L band far more important than I/Q transmission. Such as it has been mentioned in Section 4.1.2, the adequacy between, as one example, the chosen MODCOD and the satellite channel conditions, may not be granted since losses, delay and jitter in the fronthaul network could occur.

The idea is transmitting analog L band signal on optical networks is justified by its usage in commercial products, such as Cable TV (CATV) where this concept is exploited to broadcast analog video signals to subscribers. The key concept is to avoid the complexity of digitizing the L-band, a complex operation, by transmitting directly the analog signal over a fiber. It is considered as more simple to setup, as well as being more reliable and predictable. Because the signal is analog, it needs a dedicated fiber to be transmitted, and cannot be conveyed along digital signals. Hence, the capacity of the fronthaul link is not the dimensioning parameter here: either a dark fiber is available for the fronthaul link, or none is available and a dedicated network has to be built. Dimensioning this solution lies with the maximum reachable distance: because the signal is analog, the integrity of the information it conveys will be degraded by attenuation and non-linearity in the transmission.

However, the constraints inherent to the analog solution seem to overwhelm the sole performance advantages, except considering short distances for the fronthaul link, and again the availability of dark fiber. Given the expected distance of the gateways required to provide diversity or to guarantee that feeder links of the gateways do not interfer, this solution is not compatible. Moreover, the flexibility brought by capacity leasing in the digital case is a valuable asset compared to the fixed capacity of dark fiber leasing.

4.2. Synthesis on the functional isolation

In the variant A, only the network functions are centralized. Even though no functional issue could be exhibited, this variant does not fully exploit the possibilities offered by the virtualization concept, which requires firstly an isolation of the functions.

The variant B does not only centralize the network function, but also the access gateway. In the future, this may even result in the virtualization of all these functions. As opposed to the variant A, the variant B would then let more room for the virtualization of some processes. Our analysis however showed that even if it can be separated from the physical gateway, the potentially virtualized access gateway would need to be close to the physical gateway, or the specific processes that need some interactions between those gateways should be adapted. The analysis of the feasibility to centralize the whole gateway (network functions, access and physical gateways) showed that the variant C could be envisioned only if the L-band signal is digitized. The requirements for this variant would impose much more bandwidth than the variant B.

In the light of our analysis, an interesting trade-off for the centralization of the gateway, in terms of performance, cost effectiveness and feasibility, would be to isolate the network functions and the access gateway from the hub, where the physical gateway shall remain. Also, in this case we recommend to let the access gateway close to the hub, or to adapt its exchanges with the physical gateway. The processes of the network functions and access gateway are subject to virtualization.

5. VIRTUALIZATION OF THE FUNCTIONS CONSIDERING THE SATCLOUDRAN TOPOLOGY

In this section, we analyze the virtualization of satellite gateway which allows an operator to run multiple instances of virtual gateway so that each pool of virtual gateways can be used by particular SNO or assigned to different SVNOs. We focus on the variant B, where both the network connectivity and the access gateway are subject for virtualization, such as concluded in Section 4.2.

5.1. Architecture for Virtualizing Satellite Gateway functions

5.1.1. General architecture In Figure 14, we propose a general architecture of the virtualized environment.

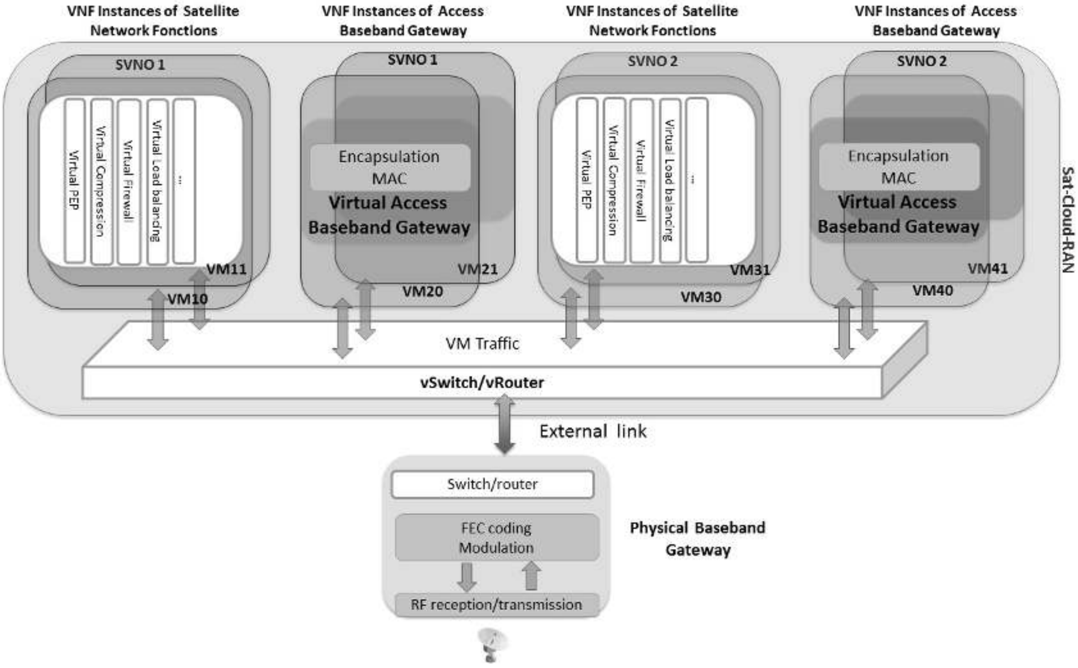

Figure 14. Virtualization of network functions and access baseband gateway

5.1.2. Virtualization environment for the network connectivity Satellite network functions which are part of the Network Gateway can be virtualized as virtual instances running in software modules described using VNFs approach. Each VNF is executed on dedicated or shared Virtual Machine (VM) so that multiple SNOs or SVNOs can be hosted on single physical infrastructure characterized by particular computational, storage and networking hardware resources.

Figure 14 illustrates the environment in which the virtual Switch (vSwitch) is responsible for switching network traffic both internally between the VMs and externally with the physical gateway. The vSwitch runs on the same server platform as the VNFs. The virtual satellite network functions such as virtual PEP (vPEP), virtual VPN (vVPN) and virtual Load Balancing (vLB) are further being connected and chained to provide a dedicated service through different VNF combinations.

For example, a first group of users of $SVNO_1$ will be provided only virtual Firewall (vFW) services as Service Function Chaining (SFC) whereas a second group will be provided vFW and vPEP service function chain.

5.1.3. Virtualization environment for the the access gateway In this scenario, the support of multiple tenants can be done using sharing or not the physical baseband gateway. We detail in Table II some use-cases on how the virtualized environment can be shared among SVNOs.

Table II. Discussion on the support of multiple tenants

Case	**Shared virtual environment**	**Shared physical gateway**	**Shared ODU**
C1		✓	✓
C2	✓	✓	✓
C3	✓		✓

In the case *C1*, each SVNO has its own virtualized environment, with dedicated pool of virtual access baseband and dedicated network gateway. The sharing is done at the physical gateway. Sharing of physical gateway means sharing of spectrum or sharing of hub sites to allow each SVNO to bring its own equipment. When the spectrum is shared, the isolation of SVNO can be realized logically at bandwidth group level. In the case *C2*, this case is similar to the previous one, except that the virtualized environment is shared between SVNO. In the case *C3*, when the physical gateway is dedicated, each SVNO bring its own modules for modulation and demodulation of traffic if the same site is used. The ODU can be shared.

5.2. Functional architecture of the SatCloudRAN

Figure 15 describes a functional architecture of the SatCloudRAN by focusing on some keys of the functionalities related to access baseband gateway, network functionalities, backhauling to the Internet, and fronthauling link to connect the instances of virtual access baseband with the physical baseband gateway. This architecture can be implemented as part of an Network Function Virtualization Infrastructure Point of Presence (NFVI-PoP) that hosts the deployed functions.

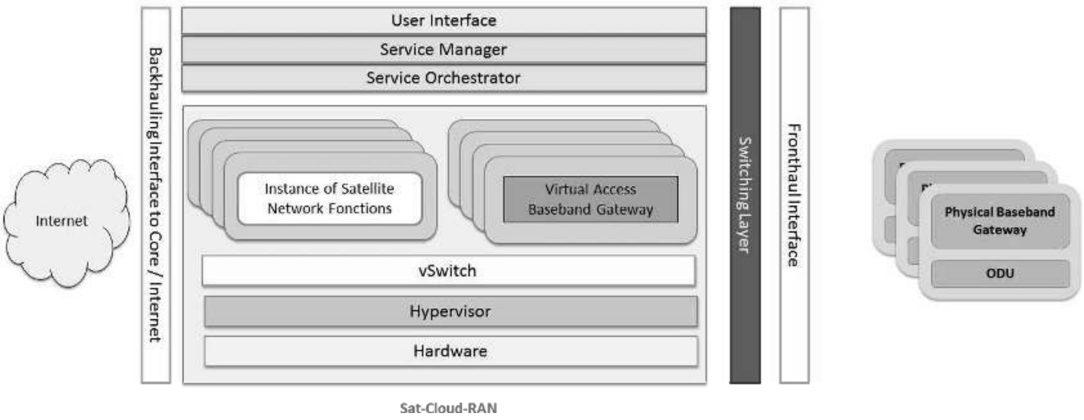

Figure 15. Overview of SatCloudRAN key functionalities installed on NFVI PoP

The design of SatCloudRAN architecture will be supported by a set of following functional elements: (1) the User Interface, that allows the operator to interact in user-friendly manner with SatCloudRAN in order to create instance and manage their functionalities. (2) The Service Manager (SM), that provides supporting services for the user interface. It interacts with service orchestrator and is responsible for managing the service orchestrator of a particular tenant. (3) The NFV manager that is in charge of the lifecycle of running VNF instance, to create, configure, orchestrate and manage the instances of created functions. The Service Orchestrator (SO) is also in charge of making in decision that needs to maintain the performance guarantee such as workload on the VMs. If any instance of function has to be scaled up or down, the SO will add or remove virtual machines and instance new instance or delete old instance to deal with the current load. Configuration will be then triggered to chain those new instances. (4) The Virtualized Infrastructure Manager (VIM) (not shown in the figure), that provides the interfaces as northbound and southbound control planes used by the SM and SO for abstracting the physical resources and instances running on cloud. (5) The service catalogue (not shown in the figure) that contains a list of the available services offered by the provider.

5.3. Discussion on VNFs

Figure 16 shows how the SatCloudRAN can be instantiated on this architecture when there are two SVNOs. The virtualization paradigm makes it easier to propose slices of virtual networks such as shown in this figure. In this view, the access gateways are not centralized. By centralizing them in a pool, it would be even easier to manage the satellite gateway diversity, since there would be only one MAC layer for a given SVNO. However, the relevance of such approach is related to the deployment, the fronthaul link characteristics and the resulting performance in terms of satellite resource utilization, quality of experience, etc.

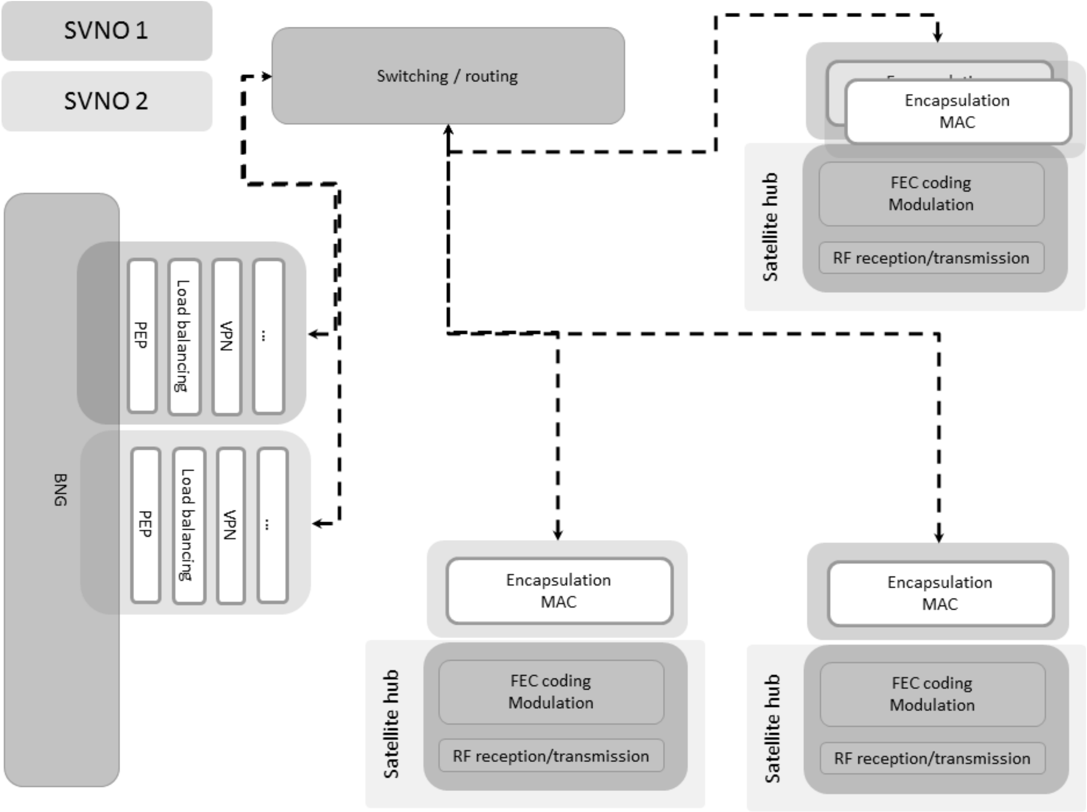

Figure 16. SatCloudRAN instanciation in a multi-gateway scenario

Since there are many interactions between an access gateway and a physical gateway, we provide some focus on them. In Figure 17, we show the interactions of the access gateway VNF with the other elements. The blue links refer to the data packets that are actually transmitted in the network (which may carry control plane information, such as the BBFRAMEs), the green links refer to the log information that could be forwarded to an SDN-controller, the red links refer to the control information exchanged between the physical gateway and the access gateway, through a physical gateway controller (the Satellite Baseband Gateway Physical Network Function (SBG-PNF) Controller).

This section will present and detail these interactions; this contribution is essential to further describe this VNF. Then the specific algorithmic elements within the acess gateway VNF can hardly be provided since it is specific to the implementation of the normative documents, but the description of these interfaces would let one Satellite Communication (SATCOM) manufacturer to propose an access gateway VNF that can be easily integrated with any existing systems, if the current approach is respected.

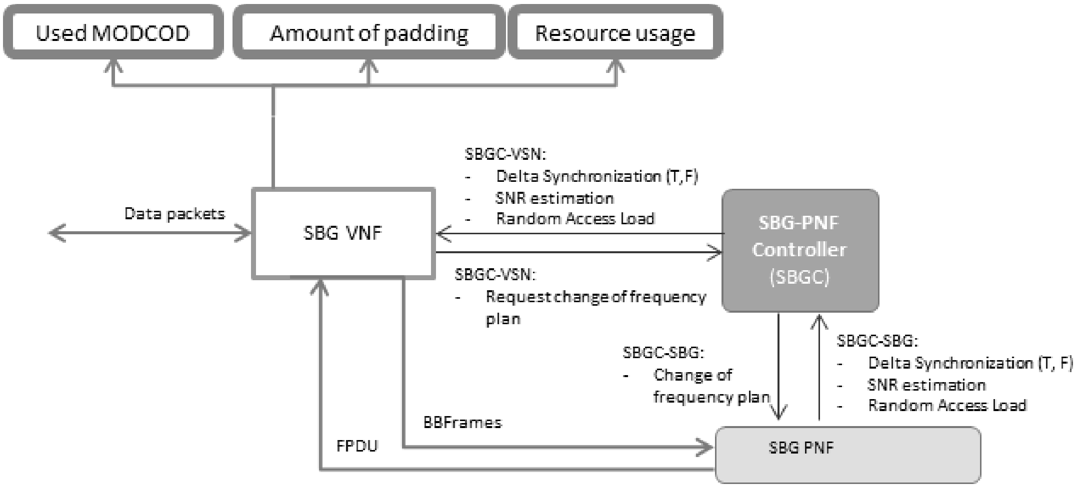

Figure 17. Input and output for the access gateway VNF

The north bound of the access gateway VNF shows the log information that can be forwarded, and taken into account in higher layer algorithms, particularly related to resource management. We propose the used MODCOD, the amount of padding and the resource usage.

The west bound of the access gateway VNF represents the data that is forwarded from/to the network gateway through the transport network.

The east bound of the access gateway VNF forwards the frequency plan changes, so that the bandwidth for the return and forward links can be adapted. This is one of the responsibilities of the RRM process.

The east bound of the access gateway VNF receives the deltas related to the synchronization in both time and frequency (related to the synchronization process), the SNR estimation on the return link (related to the FMT process) and the load measured on the random access slots (related to the RRM process). This information is measured directly on the received signal; thus it is not carried out naturally by the Frame Protocol Data Unit (FPDU). The existing SBG-PNF interface can be exploited to carry out this information, or another control interface has to be defined. The south bound of the access gateway VNF forwards the BBFRAMEs to the SBG-PNF, along with the MODCOD to apply on the BBFRAMEs. BBFRAMEs contain more information than just data plane packets, since control information and data plane packets are multiplexed in BBFRAMEs. According to the Digital Video Broadcasting (DVB) public documents, BBFRAMEs also contain Layer 2 control information, by example, information related to the way the channel capacity is shared on the return link (related to the RRM), the NCR or other information related to the network management.

The south bound of the access gateway VNF receives the FPDU from terminals. These packets contain more information than just data plane packets. As one example, they contain information related to the logon information from a given terminal (related to the logon), the resource access requests from the terminal (related to the RRM process) or the SNR on the forward link estimated by the terminal (used by the FMT algorithm). This view could be discussed since there are no specific requirements on these aspects in the public documents; however we think this approach provides a good trade-off between complexity and flexibility. The proposed approach let the access gateway VNF hosts most of the decisions related to the control processes, while the SBG-PNF is limited to the management of the actual RF resources, on which the RRM maps its allocation.

6. SDN CONTROL OF A SATELLITE CORE NETWORK

In this section, we discuss the SDN control of a satellite core network with some examples of controlled functions.

6.1. Overview of SDN controlers

SDN is envisioned to be a key enabler of the 5G to fulfill the objectives in flexibility and network programmability. This concept breaks the vertical network integration by separating the network's control logic from the underlying routers and swtiches that forward the traffic. Moreover, with a separation of the control and data planes, network switches become simple forwarding devices and the control logic is implemented in a logically centralized controller, simplifying policy enforcement and network reconfiguration and evolution.

However, it is essential to point out that network programmability is not something new, and SDN is not the only technique that could provide such flexibility and programmability. Indeed, the authors of [30] provide an historic perspective of programmable networks and clearly position the emerging concept of SDN.

The OpenFlow protocol [31] is an enabler of SDN and is being promoted by the Open Networking Foundation (ONF) [32], on the industry side, and by the OpenFlow Network Research Center (ONRC) [33] at the academic side. OpenFlow aims at standardizing the exchanges of information between the centralized SDN-controller and the components of the network. Other programmable networking efforts can be noticed and should not be neglected. However, the ONF has been able to largely gather academics, researchers and industry: this may result in OpenFlow being a de-facto standard.

The authors of [30] also propose a list of the switches and controllers that are complaint with the OpenFlow standard. In [34] assesses the maturity of five state-of-the art SDN-controllers by evaluating their capacity to process small packets based on a global view of the network. They conclude that it is necessary to rethink current SDN controllers to better leverage the energy efficiency and high network traffic capabilities. Depending on the deployment use-case, the adequacy of the SDN controllers may be questioned. That being said, specific environments such as data-centers, enterprise networks or home and small business, already exhibit the interest for including the SDN paradigm.

6.2. SDN-based control architecture

Figure 18 details the architecture for the SDN control of the virtualized environment. It is composed of a high level controller in charge of controlling and managing the entire network resources whereas low-level controller in charge of controlling and managing specific network element or domain-specific resources. For example, the Host Network Operator (HNO) controller can be used as entire network controller and SVNO controller for each SVNO subnetwork. The HNO would be in interaction with the SCC and the mission segment to evaluate the available capacities of the satellite or to update its configuration.

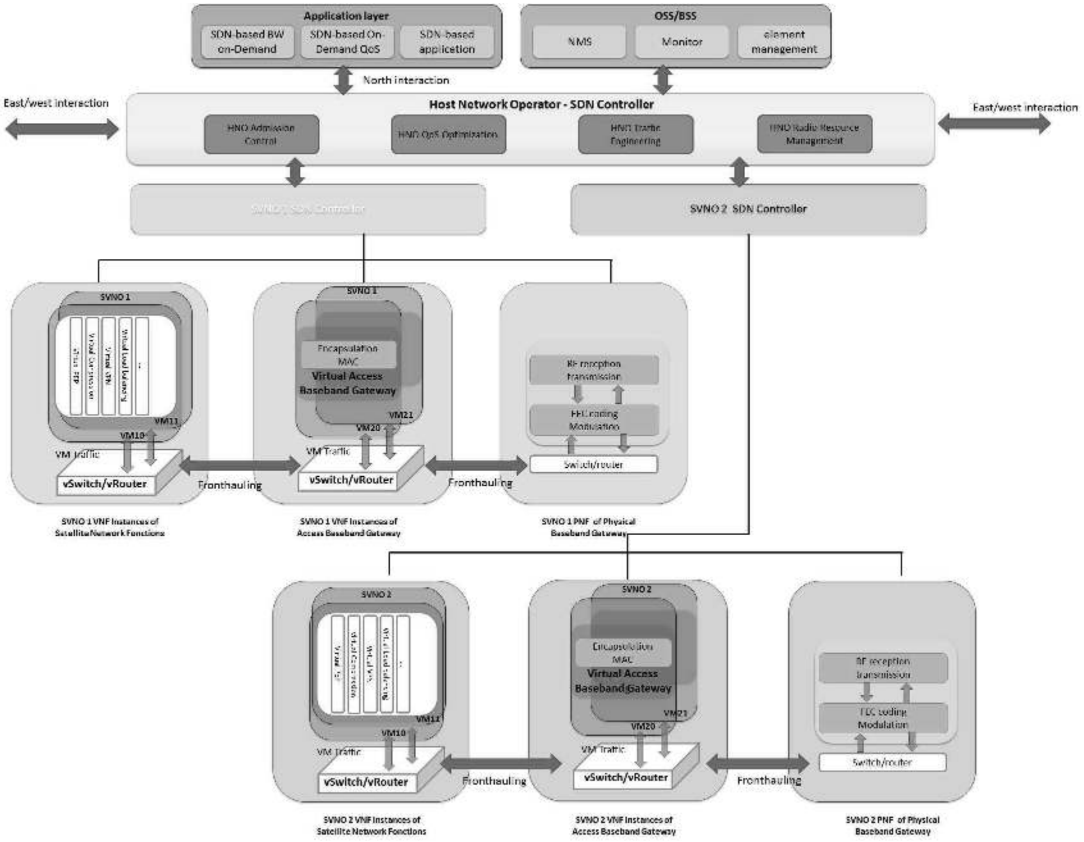

Figure 18. Multiple SDN-based control architecture for SVNO support

The SDN controller would be in charge of multiple inter-dependant modules. Firstly, the admission control is similar to call admission-control in telephony networks. The SDN controller is responsible for accepting new demand of bandwidth allocation. It allows also verifying if the network can handle the traffic demands of the application without impacting other applications adversely. Toward this end, this module needs to have an accurate view of network resources in use. Secondly, the QoS optimization is responsible for mapping (using cross-layer optimization) the traffic across different layer of networking taking into consideration the actual link performances and the defined class of services in terms of traffic classification, marking and flow control. Thirdly, the traffic engineering is responsible for centralized traffic engineering to dynamically reallocate bandwidth among different customers in case of outage or failure. It is also responsible for managing available network capacity according to application priority. Finally, the radio resources management is responsible for managing the radio resource in terms of bandwidth allocation, packet scheduling, fading mitigation technique, and how to utilize efficiently the satellite resources.

We detail in the rest of this section, three SDN-based applications that were identified in [35], namely (1) SDN-based bandwidth on demand (§ 6.3), (2) SDN-based dynamic QoS (§ 6.4) and (3) SDN-based satellite gateway diversity (§ 6.6).

6.3. SDN-based bandwidth on demand

The aim of the SDN-based flexible satellite Bandwidth on Demand (BoD) is to improve the typical satellite broadband access service with the ability to allow service providers to dynamically request and acquire bandwidth in a flexible manner. On-demand bandwidth services are established by the customer requesting change of the allocated bandwidth by interacting with corresponding SDN-based application (*i.e.* SDN-based bandwidth on-demand). This latter interacts with Admission Control (AC) function at the SDN controller using the SDN northbound Application Programming Interface (API). One of the goals of the AC function is to accept or to reject user requests of network bandwidth corresponding to a Class of Service (CoS) according to resources availability and customer SLA. Requested bandwidth parameters can be communicated to the SDN-based application through a portal or based on application and service profiles.

Indeed, the TDM forward link, which is managed by the SVNO in our architecture, is shared between customers in regard of their SLA. Bandwidth profiles are specified in SLAs to quantify agreed limits on service frame bandwidth and, as a consequence, they define traffic management operations within networks, such as policing, shaping and scheduling.

Bandwidth profiles are specified in SLAs to quantify agreed limits on service frame bandwidth and, as a consequence, they define traffic management operations within networks, such as policing, shaping and scheduling. All these tasks are managed by a bandwidth management function that may reside in the SDN controller. It allows to structure the bandwidth for groups (group of remotes, SVNO, users) an allocated this bandwidth accordingly. For each bandwidth group, it is possible to specify the bandwidth size, the priority (high or low priorities), the policy applied to the group in case of overflow or the smoothing function to adopt in case of slice overflow.

For the BoD service, it is essential to dynamically update the bandwidth profile for L2 connectivity and/or the **PHB!** (**PHB!**) for L3 connectivity. It should be noted that BoD service shall be configured to be used when needed or for a specified scheduled time. The resources released should become immediately available for other connections and usage.

6.4. SDN-based dynamic QoS

The aim of the dynamic QoS is to dynamically adjust the network connectivity characteristics and the amount of bandwidth associated with various CoS. That would result in a better match between the application needs in terms of QoS and take into account the physical characteristics. The SDN-controller would be responsible in setting the dynamic parameters and take decisions on how to update them to improve the Quality of Experience (QoE) at the end user level.

The SDN-controller must be informed about the actual CoS bandwidth capacity available for a user at a given time, the new entering flows and the conditions of fading for each terminal destination to adapt in near real time the QoS of the network and optimizing the performance of a network by dynamically analyzing, predicting, and regulating the behavior of data transmitted over that network (*i.e.* traffic engineering).

Such SDN-controller mechanisms will close the loop between the applications, network connectivity and radio resource management with the goal of providing dynamic QoS and Traffic

Engineering (TE) of bandwidth capacity for each CoS sharing the same link, with the aim to optimize end user quality of experience at all times.

For example, at L2 connectivity, if an OpenFlow switch of the SVNO receives a packet which has been never seen before and for which it has no matching flow entries to the controller, the controller can call upon the process of AC and QoS optimization to take into account this new traffic and make a decision on how to handle this flow. To help in the identification and the classification of the new packets, the controller could use the service of **DPI!** (**DPI!**) of the traffic to provide a granular control of the data flows to the QoS optimisation service in order to dynamically adapt the resources in function of the needs.

To address the dynamic QoS challenges, this SDN application could dynamically configure QoS parameters, to ensure a high QoE, based on the information of actual and predictive traffic loads and available bandwidth limited by the fading.

6.5. SDN-based satellite gateway diversity

The aim of this application is to provide a diversity scheme in the forward link by allowing multiple gateways to feed simultaneously the satellite to accommodate a high capacity aggregation using a high number of beams while ensuring gateways' resiliency. The SDN application would thus collect information related to meteorological conditions, to the detection of failure issues or to the efficiency of the resource utilization to potentially change the usage of the available gateways.

This satellite gateway diversity implies: (1) inter-gateway handover technique to cope with the cases where gateway feeder links experience outage due to meteorological conditions or failure (the handover typically implies that additional traffic is addressed towards another gateway to handle the capacity reduction of the affected gateway); (2) permanent monitoring of hub and radio resources to detect the outage, failure or any problem; (3) reconfiguration of the network capacity, optimizing traffic engineering, routing table and forwarding elements of the core satellite network provider to support temporarily capacity changes.

6.6. Applicability of the SDN to SatCloudRAN

Depending on the deployment context, dedicated controllers may be required. Control processes related to the MAC layer are indeed closely related to the characteristics of the medium used to transmit the data. This induces that related work proposes specific controllers for satellite communications [36–38]. The dynamically controlled parameters could be at the satellite platform level, at the access gateway or at the network level. The parameters of the access gateway that could be exposed to a centralized controller are deeply related to SATCOM, whereas those of the network functions may not. This section has extended the use-cases proposed in the literature and has included the SDN controllers in the frame work of the SatCloudRAN.

The deployment of the SDN paradigm within SATCOM equipments induce a split between data and control planes that is not straight forward. From a functional perspective, the logical separation of data, control and management plane functions has been actually reflected in the reference model for Broadband Satellite Multimedia (BSM) systems developed in ETSI [39]. However, it is worth pointing out that this view is not reported to be widely implemented in current systems and interoperable. If some algorithms can be considered as quite straightforward, such as the ACM operation (which is only one part of the FMT), RRM and QoS can encompass numerous complex parameters, which are not part of the standard. The interest of further splitting the access gateway in multiple VNFs shall be looked at by the satellite ground segment manufacturer since it is closely related to the specific implementation of the public normative documents.

We believe that the view proposed in this article is a workflow towards decoupling the data and the control planes and including both NFV and SDN paradigms in SATCOM. Indeed, the identification of the possibility to decouple and centralize processes of a satellite gateway shed a light on the potential issues and the need for rethinking the system. The identification of common interfaces between physical and access gateway would ease the decoupling of the data and control planes to further apply the SDN concepts in the aggregation network.

7. CONCLUSION

The role that satellite communications can play in the forthcoming 5G ecosystem is being revisited. This paper contributes to this vision through researching on the adoption of SDN and NFV technologies into the satellite domain. These concepts are seen as key facilitators to make satellite communications to become a constituent part well integrated within an anticipated heterogeneous 5G network architecture. With the introduction of SDN and NFV, greater flexibility is expected to be achieved by satellite network operators, in addition to the much-anticipated reduction of both operational and capital expenses in deploying and managing SDN and NFV compatible networking equipment within the satellite networks. This proposed concept, namely satellite cloud RAN, exploits cloud based infrastructure and data-center virtualization to deliver cost efficient, high level resources availability and flexible resources sharing. This concept shed light on better interaction and integration of the satellite network with terrestrial functionalities while supporting advance features such as traffic engineering and load balancing.

The decomposition of the satellite and network gateway into multiple functional elements allows for identifying three main splitting approaches where the virtualization benefits are associated with the gain obtained from the centralization of the functions and their multiple instantiations. It is worth pointing out that our approach can be applied to other satellite systems, since similar specific functions such as the FMT, the QoS or the synchronization would have to be dealt with.

The public normative documents do not provide information related to the implementation. We believe that going further in the chaining of the multiple internal satellite access gateway is not necessary to illustrate the interest of introducing the NFV paradigm within the SATCOM industry. Moreover, if some algorithms can be considered as quite straightforward, such as the ACM operation (which is only one part of the FMT), RRM and QoS both can encompass numerous complex parameters, which are not part of the standard.

It is the first step of access softwarization which can be seen as an overall transformation trend in SATCOM for designing, implementing, deploying, managing and maintaining access entity, exploiting characteristics of software such as flexibility and rapidity of design, development and deployment.

Splitting the access gateway in multiple VNFs can be seen as separating the data and control plane processes, each of them running as separated instances. Further than this split, one can envision that the data plane processes are further split (such as split between the generation of the GSE headers and the addition of padding), and that the control plane processes are further split (such as having the RRM, the QoS and the FMT running as separated instances).

As a future work, a proof-of-concept prototype will be designed, aiming at evaluating, among others, the practical application of virtualizing a given function and to determine the most promising virtualization capacities envisioned in this paper. The OpenSAND (ex-Platine) [40] will be used as a proof of concept platform to emulate the DVB-S2 and the DVB-RCS2 network including the physical link, the ODU, the physical gateway and BBFRAME handling.

ACKNOWLEDGMENTS

The authors are partly funded by the European Union under its H2020 research and innovation programme (grant agreement H2020-ICT-644843). The authors would like to thank the reviewers for their valuable feedback and suggestions to improve the quality of the paper.

References

1. Association GI. 5G Vision. *Technical report*, 5G PPP 2015. URL https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf.
2. NetWorld2020 - SatCom WG. The role of satellites in 5G. *Technical report*, Networld 2020 2014. URL www.networld2020.eu/wp-content/uploads/2014/02/SatCom-in-5G_v5.pdf.
3. Courville N, Bischl H, Lutz E, Svigelj A, Chan PM, Papapetrou E, Asorey-Cacheda R. Hybrid Satellite/Terrestrial Networks: State of the Art and Future Perspectives. *QShine 2007 Workshop: Satellite/Terrestrial Interworking*, IWSTI '07, ACM: New York, NY, USA, 2007; 1:1–1:7, doi:10.1145/1577776.1577777. URL http://doi.acm.org/10.1145/1577776.1577777.

4. Evans B, Werner M, Lutz E, Bousquet M, Corazza G, Maral G. Integration of satellite and terrestrial systems in future multimedia communications. *Wireless Communications, IEEE* Oct 2005; **12**(5):72–80, doi:10.1109/MWC.2005.1522108.
5. Ahn DS, Kim HW, Ahn J, Park DC. Integrated/hybrid satellite and terrestrial networks for satellite IMT-Advanced services. *International Journal of Satellite Communications and Networking* 2011; **29**(3):269–282, doi: 10.1002/sat.977. URL http://dx.doi.org/10.1002/sat.977.
6. Dai L, Chan V. Capacity dimensioning and routing for hybrid satellite and terrestrial systems. *Selected Areas in Communications, IEEE Journal on* Feb 2004; **22**(2):287–299, doi:10.1109/JSAC.2003.819976.
7. Border J, Kojo M, Griner J, Montenegro G, Shelby Z. Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations. RFC 3135 (Informational) Jun 2001. URL http://www.ietf.org/rfc/rfc3135.txt.
8. ETSI. Digital Vidideo Broadcasting (DVB); Second Generation DVB Interactive Satellite System (DVB-RCS2); Part 2: Lower Layers for Satellite standard . *Technical Report V1.1.1*, European Telecommunications Standards Institut August 2011.
9. Boucadair M, Jacquenet C. Software-Defined Networking: A Perspective from within a Service Provider Environment. RFC 7149 (Informational) Mar 2014. URL http://www.ietf.org/rfc/rfc7149.txt.
10. ETSI. Network Functions Virtualisation (NFV); Virtual Network Functions Architecture. *Technical Report V1.1.1*, European Telecommunications Standards Institute December 2014.
11. Ferrús R, Koumaras H, Sallent O, Agapiou G, Rasheed T, Kourtis MA, Boustie C, Gélard P, Ahmed T. SDN/NFV-enabled Satellite Communications Networks: Opportunities, Scenarios and Challenges. *Physical Communication* 2015; :–doi:http://dx.doi.org/10.1016/j.phycom.2015.10.007. URL http://www.sciencedirect.com/science/article/pii/S1874490715000543.
12. Mijumbi R, Serrat J, Gorricho J, Bouten N, De Turck F, Boutaba R. Network function virtualization: State-of-the-art and research challenges. *Communications Surveys Tutorials, IEEE* Firstquarter 2016; **18**(1):236–262, doi: 10.1109/COMST.2015.2477041.
13. ETSI. Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2). *Technical Report V1.2.1*, European Telecommunications Standards Institut August 2009.
14. EBU-UER DVB. Digital Video Broadcasting (DVB); Second Generation DVB Interactive Satellite System (DVB-RCS2); Part 1: Overview and System Level specification . *Technical Report V1.1.1*, European Telecommunications Standards Institut May 2012.
15. EBU-UER DVB. Digital Video Broadcasting (DVB); Second Generation DVB Interactive Satellite System (DVB-RCS2); Part 3: Higher Layers Satellite Specification. *Technical Report V1.1.1*, European Telecommunications Standards Institut May 2012.
16. ETSI. Digital video broadcasting (dvb); user guidelines for the second generation system for broadcasting, interactive services, news gathering and other broadband satellite applications (dvb-s2). *Technical Report V1.1.1*, European Telecommunications Standards Institut February 2005.

17. ETSI. Satellite Earth Stations and Systems (SES); Hybrid FSS satellite/terrestrial network architecture for high speed broadband access. *Technical Report V1.1.1*, European Telecommunications Standards Institute Mars 2015.
18. Andersson L, Madsen T. Provider Provisioned Virtual Private Network (VPN) Terminology. *RFC 4026*, RFC Editor March 2005.
19. Baker F, Fairhurst G. Ietf recommendations regarding active queue management. *BCP 197*, RFC Editor July 2015.
20. Babiarz J, Chan K, Baker F. Configuration Guidelines for DiffServ Service Classes. *RFC 4594*, RFC Editor August 2006.
21. Santitoro R. Metro Ethernet Services – A Technical Overview. *Technical report*, Metro Ethernet Forum 2003. URL http://www.mef.net/Assets/White_Papers/Metro-Ethernet-Services.pdf.
22. Dupe J, Chaput E, Baudoin C, Bes C, Deramecourt A, Beylot A. Optimized GSE packet scheduling over DVB-S2. *IEEE Global Communications Conference, GLOBECOM 2014, Austin, TX, USA, December 8-12, 2014*, 2014; 2856–2861, doi:10.1109/GLOCOM.2014.7037241. URL http://dx.doi.org/10.1109/GLOCOM.2014.7037241.
23. C-RAN The Road Towards Green RAN. *White paper*, China Mobile Research Institute 2013. URL http://labs.chinamobile.com/cran/wp-content/uploads/2014/06/20140613-C-RAN-WP-3.0.pdf.
24. I CL, Huang J, Duan R, Cui C, Jiang J, Li L. Recent Progress on C-RAN Centralization and Cloudification. *Access, IEEE* 2014; **2**:1030–1039, doi:10.1109/ACCESS.2014.2351411.
25. Fronthaul Challenges and Opportunities. *Presentation at lte world summit*, Orange Labs Network 2014. URL http://www.e-blink.com/sites/default/files/documents/LTE_world_summit-Orange.pdf.
26. Checko A, Christiansen H, Yan Y, Scolari L, Kardaras G, Berger M, Dittmann L. Cloud RAN for Mobile Networks - A Technology Overview. *Communications Surveys Tutorials, IEEE* Firstquarter 2015; **17**(1):405–426, doi:10.1109/COMST.2014.2355255.
27. C-RAN Vendors Ready for Virtualization as Asian Operators Pursue Wide-scale Deployments. *Web press release*, ABIresearch 2014. URL https://www.abiresearch.com/press/c-ran-vendors-ready-for-virtualization-as-asian-op/.
28. Virtualization of Small Cells: Overview. *Technical report 106.05.1.01*, Small Cell Forum 2015. URL http://scf.io/en/documents/106_-_Virtualization_for_small_cells_Overview.php.
29. HTS Business Models. *Technical report*, iDirect. URL http://www.idirect.net/Company/Resource-Center/Collateral-Library/~/media/Files/Infographics/HTS-Business-Models-Infographic.ashx.
30. Astuto BN, Mendonça M, Nguyen XN, Obraczka K, Turletti T. A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks. *Communications Surveys and Tutorials, IEEE Communications Society* 2014; **16**(3):1617 – 1634, doi:10.1109/SURV.2014.012214.00180. URL https://hal.inria.fr/hal-00825087, accepted in IEEE Communications Surveys & Tutorials.
31. McKeown N, Anderson T, Balakrishnan H, Parulkar G, Peterson L, Rexford J, Shenker S, Turner J. Openflow: Enabling innovation in campus networks. *SIGCOMM Comput. Commun. Rev.* Mar 2008; **38**(2):69–74, doi:10.1145/1355734.1355746. URL http://doi.acm.org/10.1145/1355734.1355746.
32. Open Networking Foundation. ; URL https://www.opennetworking.org/about/onf-overview.
33. Open Networking Research Center. ; URL http://onrc.stanford.edu/.

34. Stephen Mallon, Vincent Gramoli, Guillaume Jourjon. Are Today's SDN Controllers Ready for Primetime? *CoRR* 2016; **abs/1608.05140**.
35. Ferrus R, Koumaras H, Rasheed T, Agapiou G, Sallent O, Boustie C, Gélard P, Ahmed T. SDN/NFV-enabled Satellite Communications Networks: Opportunities, Scenarios and Challenges. *ELSEVIER PHYSICAL COMMUNICATION JOURNAL* January 2015; .
36. Nazari S, Du P, Gerla M, Hoffman C, Kim J, Capone A. Tackling Bufferbloat in capacity-limited networks. *IEEE MILCOM*, 2016.
37. Gopal R, Ravishankar C. Software Defined Satellite Networks. *32nd AIAA International Communications Satellite Systems Conference, SPACE Conferences and Exposition*, 2014, doi:10.2514/6.2014-4480.
38. Bertaux L, Medjiah S, Berthou P, Abdellatif S, Hakiri A, Gelard P, Planchou F, Bruyère M. Software Defined Networking and Virtualization for Broadband Satellite Networks. *IEEE Communications Magazine* Mar 2015; **53**(3):pp. 54–60. URL https://hal.archives-ouvertes.fr/hal-01107652.
39. ETSI. Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM); Services and architectures. *Technical Report V1.2.1*, European Telecommunications Standards Institute December 2007.
40. Baudoin C, Arnal F. Overview of platine emulation testbed and its utilization to support dvb-rcs/s2 evolutions. *Advanced satellite multimedia systems conference (asma) and the 11th signal processing for space communications workshop (spsc), 2010 5th*, 2010; 286–293, doi:10.1109/ASMS-SPSC.2010.5586897.