Towards Traffic Offload in converged satellite and terrestrial networks

Christian Niephaus, Gheorghita Ghinea
Brunel University
London, United Kingdom
Email: {christian.niephaus, george.ghinea}@brunel.ac.uk

Jens Moedeker
Fraunhofer FIT
Sankt Augustin, Germany
Email: jens.moedeker@fit.fraunhofer.de

***Abstract*—Whilst broadband Internet connectivity has become highly important, providing broadband connectivity nonetheless remains a considerable challenge, particularly in rural and remote regions where the deployment of optical fibers faces economical obstacles. A promising option to address this issue is that of the most recent satellite systems, capable of providing high capacities virtually everywhere. However, compared to most terrestrial systems, satellite networks have very different link and, more importantly, latency characteristics, which often render them only barely usable for delay intolerant traffic. Thus, convergence of terrestrial and satellite networks is required, so that only certain traffic flows can be offloaded onto a supplemental satellite connection. In this work, we propose a network architecture relying on modern Software Defined Network (SDN) concepts, which enable dynamic traffic offloading in a converged satellite and terrestrial network, in order to relieve the load in a narrowband terrestrial network. We show that with limited overhead, a traffic can be offloaded, leading to an increase in the user's Quality-of-Experience (QoE).**

I. Introduction

Broadband Internet connectivity has become increasingly important over the past years and is nowadays considered "a crucial factor to realize economic growth". This enables the development of new services and applications, many video-based [1], [2], as described in the European Digital Agenda [3], [4] Hence, in this agenda the European Commission sets the objective to enable broadband Internet connections of at least 30 Mbps to be available to all EU citizens and 100 Mbps to at least half of European households by 2020. Moreover, it is widely accepted that current and emerging networks need to be able to cope with a tremendous increase of traffic volume over the next years, as predicted in [5].

This significantly higher amount of traffic volume will pose a major challenge for operators, especially in rural or other difficult-to-serve areas. For example, in the backhaul segment, the deployment of nowadays' typically-used technologies, such as optical fibers, microwave radio links or copper connections [6] is prevented by economical constraints [7] leading to many underserved areas.

In order to address this issue, a promising approach is to rely on satellite systems. Bidirectional satellite networks recently regained the attention of both the scientific and industry communities since the next generation of Geostationary Earth Orbit (GEO) fixed satellite systems, which are scheduled to be operational by 2020, will be targeting the Terabit/s aggregated capacity [8]. These systems will lower the cost per bit significantly mainly by transmitting on different frequencies, i.e. Ka-band, and by using multiple but relatively small spots. Those spots have a size of a few hundred kilometers (instead of e.g a single spot for the whole of Europe) and, thus, allow for more spatial frequency re-use [9]. As satellite links provide ubiquitous and resilient services as well as broadband coverage, they are able to deliver high throughput connections and their additional capacity wherever it is needed, on a very flexible basis.

However, compared to most terrestrial network technologies, both wired and wireless, satellite links have highly different characteristics in terms of latency, burstiness or link stability. Even though approaches exist, that aim at providing triple-play services over broadband GEO satellites, it is de facto impossible to achieve a similar QoE perceived by the users compared to broadband terrestrial networks, when real-time and interactive application are being used [10]–[12]. The higher latency on satellite connections as a result of the high signal propagation time inevitably means the user's QoE is lower compared to a broadband terrestrial link.

Thus, using solely new satellite systems to provide broadband Internet access in rural and remote areas would not solve the problem, since user expectations in terms of location-independent Quality-of-Service (QoS) and QoE, cannot be met due to the high latency. That is, a user in a remote area expects the same high quality service as a user living in an urban region. Instead, satellite networks need to be integrated as a native component into existing terrestrial infrastructures as already acknowledged previously [13]–[15]. Thereby, a converged satellite and terrestrial network is formed, so that traffic can be offloaded dynamically to the satellite connection, whenever it makes sense.

Hence, in this work, we present a network architecture relying on SDN concepts that enables dynamic offloading of traffic onto a satellite connection within a framework of integrated satellite and terrestrial networks, so that the load on the terrestrial link is reduced, while the QoE does not suffer due to the characteristics of the satellite link. Given that satellites are primed for broad- and multicast traffic, offload such traffic is particularly appealing.

The remainder of this paper is structured as follows: firstly,

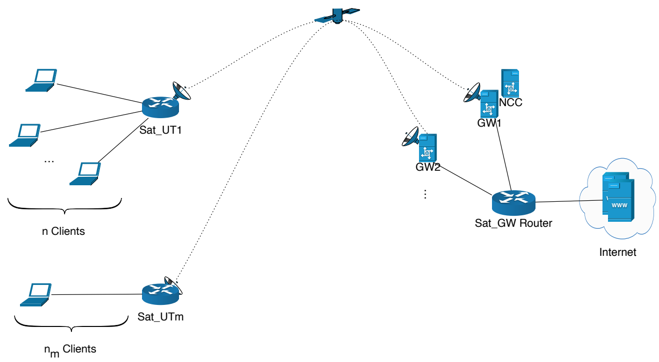


Fig. 1. Typical architecture of a satellite network

we present the related work, elaborating on satellite and terrestrial access networks, as well as the convergence thereof and typical offloading approaches. Afterwards, we describe in detail our approach, following a validation. Finally, we conclude and discuss future work.

II. Related work

In this section, we start off by presenting related information in respect of both satellite and terrestrial networks. Afterwards, we elaborate on convergence of both. In the second part we then focus on different offloading techniques already used in other networks architectures.

A. Satellite network part

The typical architecture of a satellite network is depicted in Fig. 1. It usually consist of one or multiple so-called Gateways (GWs) or hub stations, which provide on the one hand connection to other networks, i.e. the Internet, and on the other hand establish the link to the satellite. The GWs are interconnected by high speed Internet Protocol (IP) networks. The so-called User Terminals (UTs) form the other edge of the satellite network. They receive the downlink data from the satellite but also send back the uplink packets. The end user clients, such as Laptops or Smartphones, are connected by Local Area Network (LAN) technologies, e.g. Ethernet or Wireless Local Area Network (WLAN) to the UTs. Usually co-located with one of the GWs is the network control center (NCC), which has the responsibility to assign time slots to the UTs when they can send data. This is only required for the uplink, as the UTs are multiple distributed senders, which compete for the same transmission slots.

In this work, multi-beam bi-directional GEO satellite connections based on Digital Video Broadcasting-Satellite - Second Generation (DVB-S2) [16] and DVB Return Channel via Satellite (DVB-RCS) [17] are considered. That is, the satellite can be used for up- and down-link traffic. Moreover, even though several aspects discussed in this work are also applicable for Mobile Satellite Services (MSS), Fixed Satellite Service (FSS) networks are assumed.

We also assume that the satellite network provides a transparent IP connection. Since satellites are primarily used for broadcasting content, while access networks in general primarily transport unicast data, both worlds need to be brought together. That is, the IP domain and its concepts and mechanisms, which are usually used on access links, and the satellite domain. In order to do this, for instance DVB Generic Stream Encapsulation (DVB-GSE) [18] has been standardized by the European Telecommunications Standards Institute (ETSI). DVB-GSE provides an efficient encapsulation method for IP packets over variable length Layer 2 fragments, which are then directly scheduled on the physical layer into base-band frames and, hence, make the satellite connection transparent to IP.

It should be noted again, that a major difference between satellite and terrestrial systems is the latency. More precisely, GEO satellite links have a high fixed latency. While dynamic latency consists of the time required to serialize data, process a packet at a network entity, as well as potential queuing and buffering delays as well as the time required to get access to the medium, the fixed part of the overall latency is the actual signal propagation time that a packet experiences in any case when being transported over a network medium. It is a physical characteristic, mainly restricted by the speed of light. Moreover, the dynamic latency highly depends on the available capacity on the link and can be reduced and controlled by prioritization, packet dropping or admission control to avoid congested links. Hence, in terrestrial networks low latency and low jitter values can be ensured by just avoiding congested links and, thus, preventing extensive queuing [19]. GEO satellites operate at a height of 36.000km. Therefore, the fixed latency on a satellite link is in the order of magnitude of a few hundred milliseconds leading to a high overall latency even if sufficient capacity on the satellite link is available.

1) Terrestrial network part: The network architecture of the terrestrial network part considered in this work is an all-IP network relying on various transmission technologies. That is, on the data plane the regular IP protocol is used. It is assumed that clients are connected to a Customer premises equipment (CPE) that corresponds to the UT in the satellite network. The connection to the operator network is assumed to be realized via a tiered network architecture. That is, a last mile tier, such as X-Digital Subscriber Line (xDSL), and one or multiple aggregations networks form the backhaul part to the operator's network, where the interconnection point with other networks is located. This general architecture is depicted in Fig. 2.

It should be noted that for the reminder of this work the exact architecture of the terrestrial network are not particularly relevant, as long as it is an all-IP network. However, in order to benefit from an additional high capacity satellite connection, it is assumed that the throughput, which can achieved by the end user via the terrestrial connection is rather limited compared to the goals defined by the EU and other bodies.

B. Converged satellite and terrestrial networks

Historically satellite networks are often used as access networks in areas without any terrestrial connection, as shown

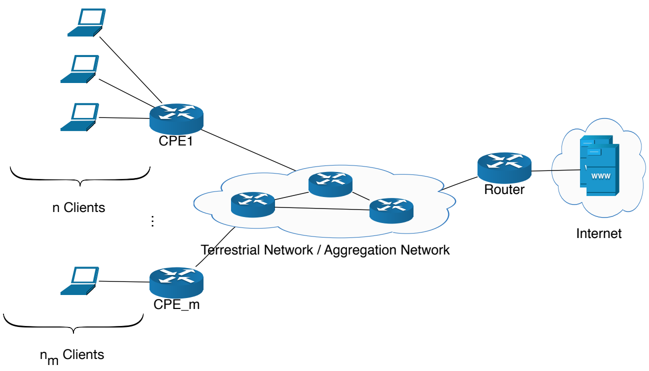
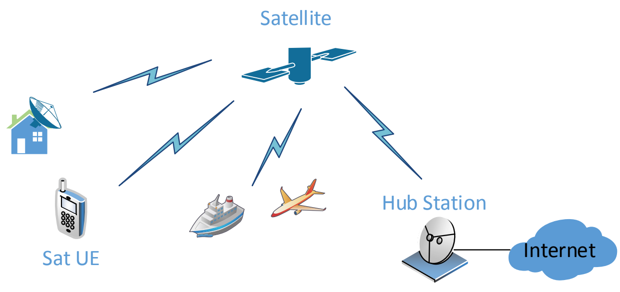
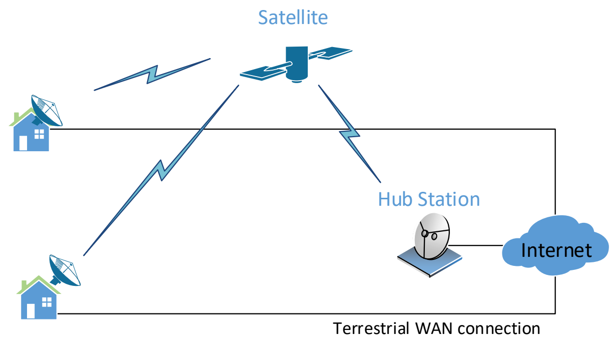


Fig. 2. Terrestrial WAN architecture



(a) Satellite network extending a terrestrial network



(b) Satellite network supplements a terrestrial network

Fig. 3. Comparison of satellite link integration options

in Fig. 3(a). These scenarios typically include providing connectivity to very remote households and premises or moving locations, such as a vessel or an airplane [20]. Typically, on one or both edges of the satellite link, terrestrial networks are connected. For instance, in a remote household typical end user devices are connected locally via a (W)LAN router connected to a satellite modem, which establishes the satellite connection or on an airplane in which the satellite connection is distributed via WLAN. However, it is important to note that the satellite link is the sole connection, so that the traffic cannot be routed differently. Such a setup usually aims at providing general connectivity.

Various scenarios and use cases following this setup have already been widely discussed and surveyed in the literature along with the associated challenges, e.g. [9], [20]–[25]. Hence, in this work we focus on scenarios where the satellite network provides an additional connection in parallel to an existing terrestrial one, as shown in Fig. 3(b). A high capacity satellite network supplements an existing terrestrial connection, as presented above, in order to increase the performance of a terrestrial network, e.g. a household that is connected with a terrestrial Digital Subscriber Line (DSL) connection as well as a satellite link. Such a scenario makes sense if the terrestrial link performs insufficiently and is not able to cope with the traffic demands. More precisely, the results of [26] suggested that such a converged satellite and terrestrial scenario can play a role if the terrestrial fixed line speed for a household is between 0-8Mbit/s. In contrast to the previous setup, the satellite link provides not the sole connection but an alternative one. Thus, traffic can be distributed onto both the terrestrial and the satellite network.

We presented a more detailed view of the challenges and the current state of the art regarding converged satellite and terrestrial networks in [27].

C. *Traffic offloading*

Offloading certain traffic from a primary network onto other networks that are more suitable becomes appealing in many scenarios. Particularly in mobile networks, which are struggling, with the increasing amount of traffic, offloading is seen as one method to mitigate capacity constraints [28]. Thus, data offloading techniques have been specified recently by 3rd Generation Partnership Project (3GPP). The concept behind data offloading is to utilize other available networks, such as a local WLAN, and send parts of the traffic via this connection instead of the actual cellular network.

In order to enable offloading in mobile networks, three different key offloading techniques exist, which are Selected IP Traffic Offload (SIPTO), Local IP Access (LIPA) and IP Flow Mobility (IFOM), as described in [29]. LIPA aims at route traffic to private networks directly instead of the regular data path through the evolved Packet Core (EPC). For this purpose, it requires a local femto cell, i.e. a Home eNB (HeNB), which acts as the so-called break-out point. In contrast, when IFOM is used, it is the responsibility of the Mobile Terminal (MT) to establish multiple connections via the available networks, e.g. an additional connection via WLAN besides the regular cellular connection, and to distribute the traffic onto these connections. Both are not applicable to the scenario considered in this work, since we neither assume the presence of HeNBs nor that the end user devices have the capabilities to communicate directly with the satellite.

Furthermore, SIPTO does not provide an adequate solution to the aforementioned challenges, since the basic idea behind it is to select data gateways, to which the MTs connect to, geographically and in terms of network topology as close as possible the Radio Access Network (RAN). In an extreme case those gateways might be even located at the Evolved Node B (eNodeB) itself, so that the traffic can be redirect without being forwarded to the EPC. Moreover, SIPTO often uses Access Point Names (APNs) to differentiate which portion of the traffic should be offloaded, which also does not provide the required level of traffic differentiation. A holistic overview of offloading techniques in mobile networks can be found in [28].

One of the major challenges with respect to traffic offloading is the selection of traffic that should be offloaded. While it is fairly easy if all traffic of a certain user or a device should be offloaded, it becomes difficult if only a certain portion should be routed differently. For example, the simplest so-called on-the-spot offloading commonly used on mobile devices offloads all traffic to WLAN as soon as a WLAN becomes available [30]. Due to the specific characteristics of the satellite network this is not suitable in this scenario.

Likewise, approaches that are used in data centers environments, such as [31], where an OpenFlow-based framework is used to realize dynamic load balancing scheme for data- center networks, cannot be easily adopted for integrated satellite and terrestrial networks. In [31], the authors prose to treat so-called elephant flows (long-lived and high traffic) differently compared to mice flows (short-lived and low traffic), in order to achieve a benefit compared to typical Equal Cost Multipath (ECMP) routing. However, even though this approach already routes traffic based on its characteristics, it is not sufficient for the specific scenario considered here, due to the high latency that needs to be taken into account on satellite links.

III. Proposed Overall Architecture

In this section, we first present our view on the integration of satellite networks into existing terrestrial network infrastructures. Afterwards, we discuss our approach on enabling offloading of traffic onto the satellite connection in such a network architecture.

A. Integration of satellite networks

Satellite and terrestrial networks can be integrated in various ways. However, the main challenge that needs to be addressed by the network architecture of such networks is always to perform proper routing decisions in order to select the optimal link for certain traffic, given the traffic's characteristics as well as the current load situation in the network. The integration itself can be very loose, i.e. both networks exist and operate virtually independently of each other and the end user system simply picks one of the two network for a certain portion of traffic. Hence, each end user device needs to be able to handle two networks and distribute the traffic. This, however, would require modifications on each end user system in order to select one of each networks in a beneficial manner. Thus, in this work an architecture for a converged satellite and terrestrial network is assumed to have additional devices which interconnect both networks.

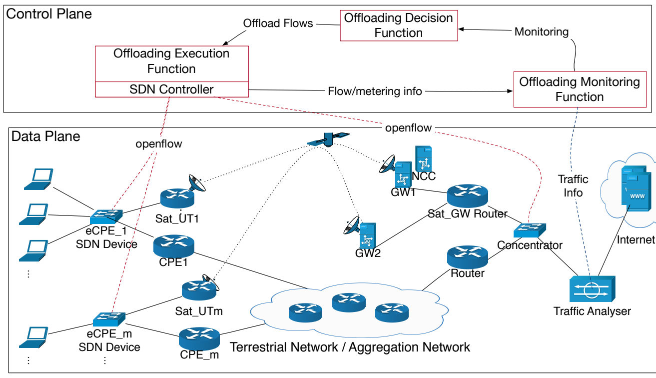


Fig. 4. Converged network architecture

These are the so-called extended Customer premises equipment (eCPE) devices as well as a Concentrator device located in the Data Plane, as depicted in Fig. 4. While the end user devices connect to the eCPEs in order to have access to both network, the concentrator bridges the satellite and the terrestrial network in the operator's network part. If a single operator does not own both networks, this will be most likely the terrestiral operator as it is assumed that the default path between the end user clients and the operator's network and, thus, the Internet as such is the terrestrial connection, since it has a higher responsiveness as long as it is not over utilized. Moreover, given that deep knowledge on the traffic itself is required in order to identify the traffic that is best suited to be offloaded to the satellite, a traffic analyzer can be used to perform e.g. Deep Packet Inspection (DPI) analysis. Due to this architecture, no modification or awareness on the end user systems is required, as eCPEs and the concentrator are the locations where the decision as to which flow is sent via which path, is executed. We acknowledge that the concentrator is a single point of failure in the network, but we believe that this risk can be strongly mitigated by using modern high availability techniques.

Furthermore, given the expectation that traffic flows need to be distributed very dynamically between the terrestrial and satellite part, relying on emerging SDN techniques seems a promising approach as e.g. explained in [32]. Hence, in the proposed architecture, both eCPEs and the concentrator are SDN-enabled network devices, i.e. OpenFlow switches. Moreover, relying on SDN creates inevitably a dedicated control plane, as depicted in Fig. 4, where the required SDN controller is located. The controller uses the OpenFlow protocol to configure the OpenFlow devices, i.e. all eCPEs and concentrator devices. Moreover, besides the SDN controller, additional components are located in the control plane, which are required to allow traffic offloading, as explained in the following section. These components can be considered as network applications in the SDN terminology.

B. Offloading architecture

In order to enable offloading of traffic from the terrestrial connection, the packet flow between end user clients and the operator's network needs to be changed from the default terrestrial path towards the satellite path for certain traffic flows. That is, the packet flow within the concentrator and/or eCPE needs to be dynamically modified, so that the alternative route is used by the packets of the specific flow. Given the specific characteristics of the satellite link, offloading makes particularly sense for videos flows and delay tolerant flows that transport many data. However, by just integrating SDN devices as well as a SDN controller, traffic offloading is not enabled per se, since on the one hand the decision to offload a particular flow needs to be made by some entity, and on the other hand, the decision needs to be transformed into specific SDN-flow rules, which can be implemented into the flow tables of the SDN devices. Moreover, in order to be able to make an informed decision monitoring information is required.

Therefore, as described previously, we propose a control plane that adapts the typical SDN architectures [33], but enriches them with several additional components.

First, a Offloading Decision Function (ODF) that is responsible for making a decision to offload a certain flow based on various information, such as load in the network or traffic characteristics. Due to the centralized SDN approach, the ODF can be provided with a global view on the network and, thus, modern approaches e.g. reyling on machine learning can be applied. It should be noted, that the actual decision process within the ODF is outside of the scope of this paper.

The ODF communicates its decision to an Offloading Execution Function (OEF) that is responsible for executing it. That is, the OEF receives a 5-tuple flow description of one or multiple flows, potentially including some wild-cards, that should be offloaded to the satellite link. Thanks to the global view, the OEF identifies the devices on which the flow rules needs to be changed in order to offload the flow. Moreover, the OEF generates proper OpenFlow rules, which are sent to the SDN controller, so that they can be forwarded to the devices using the regular OpenFlow protocol. More precisely, the created rules instructs eCPE and concentrator to modify the layer 2 destination address as well as the outgoing port of the packets belonging to the flow, which should be offloaded, so that packets are sent towards the satellite link.

Moreover, in order to perform a more optimal offloading decision the SDN controller feeds back monitoring information, such as the appearance of a new flow or other metering values, to a Offloading Monitoring Function (OMF). The information is not provided directly to the ODF in order to allow for enriching it with additional knowledge gained directly from network components, such as load on the links or available capacity. Particularly in satellite networks the latter might vary due to adaptive coding and modulation (ACM) techniques. Also, an additional traffic analyzer can perform e.g. DPI or other techniques to investigate the traffic in order to identify application types. Furthermore, collecting statistics on the flows from the eCPEs generates a lot of additional traffic that needs to be transmitted in-band over the Wide Area Network (WAN) connection and, thus, reduces the bandwidth that is available for transmitting user data. In order to reduce this, the OMF dynamically request this information only if needed, since e.g. in many cases the statistics available at the Concentrator are sufficient. Similar to the OEF, the OMF creates the proper OpenFlow rules to request the flow statistics from Concentrator and, if required, also from eCPEs. These OpenFlow messages are sent via the SDN controller to the devices.

The information is then provided to the ODF by the OMF. Thus, ultimately a control loop is formed that consist of a decision, execution and monitoring component which are implemented by ODF, OEF and OMF, respectively.

Furthermore, it should be noted that the location of the control plane is actually flexible, given the usage of the OpenFlow protocol, i.e. just an IP connection is required. However, in order to optimize the latency as well as to minimize the amount of generated control traffic in the WAN, the control plane should be located close to the Concentrator.

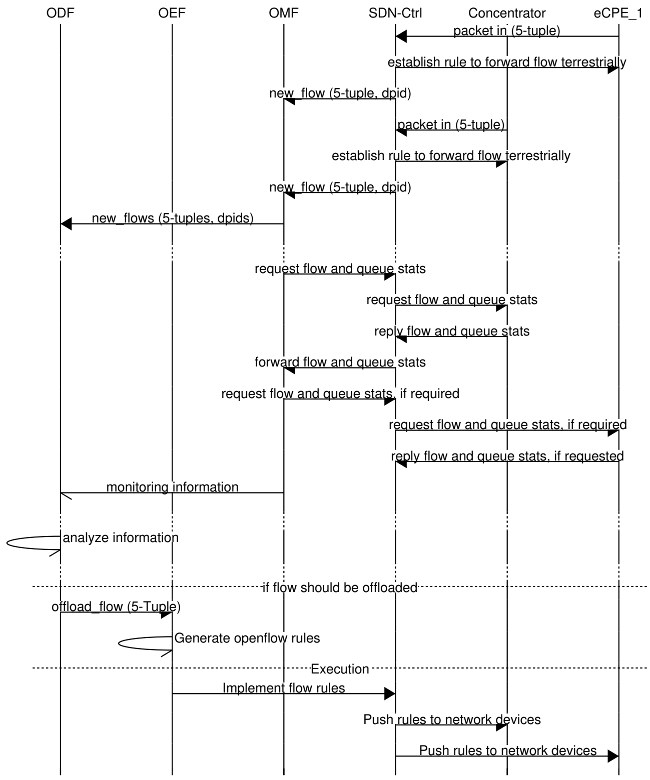


Fig. 5. Offloading sequence chart

The exact message exchange is depicted in Fig. 5. As can be seen, the concentrator and eCPE notify the SDN controller about a packet that belongs to a new flow, i.e. a new flow 5-tuple. The SDN controller then immediately instructs the notifying devices to forward the flow terrestrially, in order to avoid a long delay in the startup phase of a flow. The OMF analyzes the messages, joins them and forwards them to the ODF. Moreoever, if required by the OMF, the SDN controller also requests regularly from concentrator as well as eCPE queue and flow statistics, which are again forwarded to the ODF. In addition to that information on the traffic itself that are gained by the traffic analyzer is also forwarded via the OMF to the ODF.

The ODF in turn takes all this information into account and decides if certain flows should be offloaded. If so, it sends the corresponding 5-tuple to the OEF. Upon reception, the OEF generates proper OpenFlow rules that modifies the flow tables of the involved devices as mentioned previously to send the traffic to towards the satellite link. These rules are sent by the OEF to the SDN controller that pushes these rules onto the devices.

IV. Validation

In the following, we validate the proposed architecture to ensure that it is capable to offload traffic. We also evaluate the amount of additional control traffic it generates and as well as its impact on the Page Load Times (PLTs) of web sites.

In order to conduct these simulations, we rely on the Discrete-Event Network Simulator NS3[1], including the OF-Switch13 add-on [34] that allows for simulating an both OpenFlow 1.3 switch network device as well as a controller application interface. Moreover, to simulate the satellite the outcome of the ESA project "Development of an Open-Source, Modular and Flexible Satellite Network Simulator" called SNS3 has been used [35]. All of these components are well-known and commonly used tools in academic studies.

In order to generate network traffic, different kind of application are simulated. First of all, Hypertext Transfer Protocol (HTTP) traffic is created using the traffic generator of the SNS3 module is, as described in [36]. This module is based on well-known standards [37]. Http traffic is simulated by an on/off application. During the ON periods an object is requested by a client from a server and transferred from the server to the client. In contrast, the OFF period simulates the user reading the website. It should be noted that each call from a client first requests a main object which then might lead to further requests of so-called embedded objects. This simulates the structure of a real website which usually consist of a main webpage and several embedded objects such as images, IFrames, JavaScripts, etc.. The exact random parameter of the simulated HTTP traffic are summarized in [37]. In order to mimic typical video streaming traffic, User Datagram Protocol (UDP) flows are used. Precisely, a simulated movie using the MPEG4 codec is used. The network simulator uses a capture file that describes frames in terms of type, time and length [38] and generates traffic accordingly. Thus, the network bandwidth required to transmit the movie is not constant but varies during the transmission period. Finally, as described in [34], the delay introduced to simulate the average flow table search time is estimated as $k * \log_2(n)$, where k is the constant attribute set to the time for a single hardware operation and n represents the current number of flow entries in the pipeline. In our simulations k is configured to $20\mu s$, which is the default value for OFSwitch13.

[1]https://www.nsnam.org

Moreover, the simulation setup follows the network architecture shown in Fig. 4. The terrestrial connection between CPEs and the router in the operator edge of the network are simulated by Point-to-point links limited to a capacity of 1mbps each. However, just one eCPE is used, connecting a various amount of clients. It should also be noted that the during these simulations the ODF statically decides to offload all flows that are using the UDP protocol with a destination port between 4000 and 5000 and are seen on one eCPE as well as the concentrator, since the actual algorithm running within the ODF is out of scope in this work.

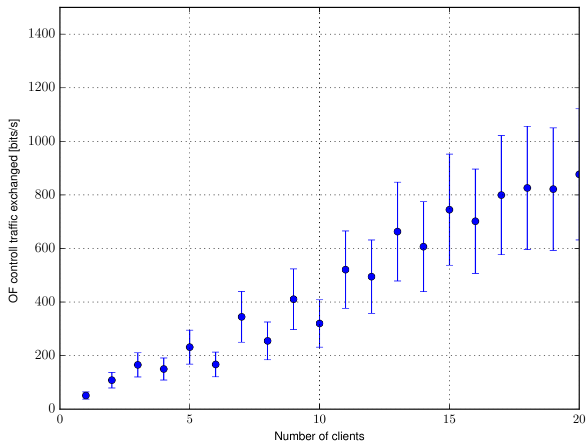


Fig. 6. Amount of control traffic and flows depending on number of eCPEs

First, we want to evaluate the amount of control traffic required to enable the traffic offloading depending on the number of clients. Thus, we run the simulation with increasing number of clients, starting from 1 client up to 20 clients. Each client creates the aforementioned simulated web traffic and video stream. The simulation runs for 180s and is repeated four times. As can be seen in Fig. 6, the amount of SDN control traffic, increases virtually linearly. Given that each new HTTP request uses a new source port, this is obvious, since every flow is considered as a new flow by the SDN controller and the ODF. Hence, each flow requires a control traffic exchange between controller and eCPE as well as controller and concentrator. It should also be noted, that the flow timeout is set to unlimited and concentrator send the first 128bytes of each packet to the controller, in case no flow rule exist. Thus, the amount of flow statistics also increases with every new flow as well as the amount of flow statistics. As show in Fig. 6, the amount of control traffic does is almost 900kbps if 20 clients are active and, thus, consumes 90% of the available bandwidth.

Hence, we test the impact on the flow timeout value on the amount of control traffic. Therefor we run the same simulation again but with a fixed amount of 10 clients. Instead of changing the amount of clients, we modify the time until an entry in the flow table of eCPEs and concentrator times out, starting from 1s up to 30s.

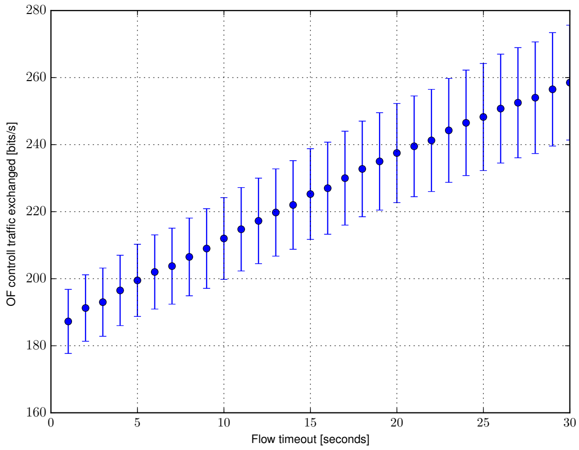


Fig. 7. Amount of control traffic and flows depending on the flow timeout

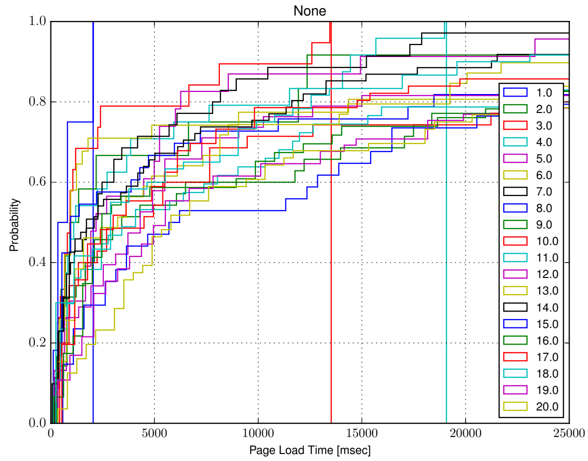
As shown in Fig. 7, with an increasing flow timeout the control traffic rate also increases linearly. This is mainly caused by the additional statistics data that need to be transmitted. Moreover, in the given traffic scenario, besides the video traffic most flows are finished before the flow timeout is reached. As can be seen, the control traffic rate differs from approx 190kbps if the flow timeout is just 1s and increases up to 260kbps for a flow timeout of 30s.

Finally, in order to further evaluate the impact of our offloading architecture, we measure the PLTs and compare them to the same network operating without our approach, since the PLT is one of the main KPIs to that impacts the QoE [39]. The results are shown in Fig. 8. As can be seen in Fig. 8(a), even with 20 clients around 90% of all websites are loaded in less than 10s and more than 82% in less than 5s. In Fig. 8(b) the results of the same test without our offloading architecture is shown. As can be seen, for 20 clients only approx. 50% of the website are loaded within 5s. Given that 10s is usually the limit for keeping the users' attention, it can be seen as critical that only between 50% and 90% of the web sites are loaded within this time frame. [40].

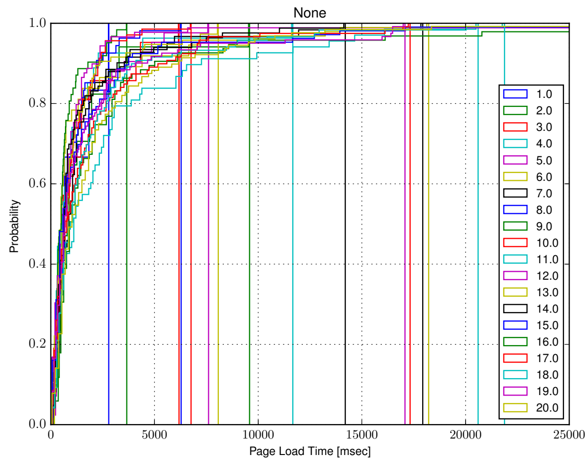
V. Conclusion and Future Work

In this work, we have presented a SDN-based converged satellite and terrestrial network architecture than allows for offloading of traffic onto the satellite link. Such networks harbor a promising option, to overcome bandwidth shortcomings in rural and remote areas, if the satellite networks with their high capacity but also high latency can be integrated successfully. The presented offloading architecture is one of the key building blocks towards this integration as it allows for selective offloading of flows onto the satellite connection. By running multiple simulations we have shown that the additional overhead introduced by our approach is virtually negligible.

Future work will mainly focus on designing and implementing ODF. In particular which information is required and most useful in order to perform a most optimal offloading decision. Moreover, enhancing the OMF and investigating the benefit of creating interactions between OMF and NCC, so that the current status of the satellite connection can be better estimated and the ODF can perform a more informed decision, is also part of the future work.



(a) With offloading architecture



(b) Without offloading architecture network

Fig. 8. Comparison of PLTs with and without offloading architectures.

References

[1] V. P. K. M and S. Mahapatra, "Quality of experience driven rate adaptation for adaptive HTTP streaming," IEEE Transactions on Broadcasting, pp. 1–19, 2018.

[2] G.-M. Muntean, P. Perry, and L. Murphy, "A comparison-based study of quality-oriented video on demand," IEEE Transactions on Broadcasting, vol. 53, no. 1, pp. 92–102, mar 2007.

[3] L. Zou, R. Trestian, and G.-M. Muntean, "E3doas: Balancing QoE and energy-saving for multi-device adaptation in future mobile wireless video delivery," IEEE Transactions on Broadcasting, vol. 64, no. 1, pp. 26–40, mar 2018.

[4] European Commission - Directorate-General for Communication Networks, Content and Technology (CONNECT), "Digital agenda for europe scoreboard 2012," June 2012.

[5] Cisco Cooperation, "The Zettabyte Era-Trends and Analysis," Whitepaper, 2013.

[6] P. Briggs, R. Chundury, and J. Olsson, "Carrier ethernet for mobile backhaul," IEEE Commun. Mag., vol. 48, no. 10, pp. 94–100, October 2010.

[7] D. Henkel, S. Englander, M. Kretschmer, and C. Niephaus, "Connecting the unconnected - economic constraints and technical requirements towards a back-haul network for rural areas," in Proc. Global Commun. Conf. Workshops, Dec 2011, pp. 1039–1044.

[8] B. Evans, P. Thompson, L. Castanet, M. Bousquet, and T. Mathiopoulos, "Concepts and technologies for a terabit/s satellite," in Proc. 3rd Int. Conf. on Advances in Satellite and Space Communications, Budapest, Hungary, 2011.

[9] A. Botta and A. Pescape, "New generation satellite broadband internet services: Should adsl and 3g worry?" in IEEE 32nd Int. Conf. Comput. Commun. Workshops, April 2013, pp. 399–404.

[10] Y. Chen, T. Farley, and N. Ye, "Qos requirements of network applications on the internet," Inf. Knowl. Syst. Manag., vol. 4, no. 1, pp. 55–76, 2004.

[11] Internet protocol aspects – Quality of service and network, ITU-T Recommendation Y.1540, 2011.

[12] Network Performance Objectives for IP-Based Services, ITU-T Recommendation Y.1541, 2003.

[13] G. Araniti, I. Bisio, M. D. Sanctis, A. Orsino, and J. Cosmas, "Multimedia content delivery for emerging 5g-satellite networks," IEEE Transactions on Broadcasting, vol. 62, no. 1, pp. 10–23, March 2016.

[14] C. Niephaus, M. Kretschmer, and G. Ghinea, "Towards qos-aware load distribution in heterogeneous networks," in Proc. IEEE 11th Malaysia Int. Conf. on Commun., Nov 2013, pp. 151–156.

[15] B. Evans, M. Werner, E. Lutz, M. Bousquet, G. Corazza, G. Maral, and R. Rumeau, "Integration of satellite and terrestrial systems in future multimedia communications," IEEE Wireless Commun., vol. 12, no. 5, pp. 72–80, Oct 2005.

[16] Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2), ETSI EN 302 307, Rev. V1.3.1, March 2013.

[17] Interaction channel for Satellite Distribution Systems, ETSI EN 301 790, Rev. V1.5.1, May 2009.

[18] Digital Video Broadcasting (DVB); Generic Stream Encapsulation (GSE); Part 1: Protocol, ETSI TS 102 606-1, Rev. V1.2.1, July 2014.

[19] X. Xiao and L. Ni, "Internet qos: a big picture," IEEE Netw., vol. 13, no. 2, pp. 8–18, Mar 1999.

[20] M. Breiling, W. Zia, Y. Sanchez de la Fuente, V. Mignone, D. Milanesio, Y. Fan, and M. Guta, "Lte backhauling over meo-satellites," in Proc. 7th Adv. Satellite Multimedia Syst Conf. and 13th Signal Process. for Space Commun. Workshop, Sept 2014, pp. 174–181.

[21] S. Kota and M. Marchese, "Quality of service for satellite ip networks: a survey," Int. J. of Satellite Commun. and Netw., vol. 21, no. 4-5, pp. 303–349, 2003.

[22] F. Zampognaro, "Satellite and terrestrial network integration," Ph.D. dissertation, University of Rome Tor Vergata, 2009.

[23] T. Gayraud and P. Berthou, "A qos architecture for dvb-rcs next generation satellite networks." EURASIP J. Wireless Commun. and Netw., vol. 2007, 2007.

[24] L. Audah, Z. Sun, and H. Cruickshank, "Qos evaluation of multiservice applications over integrated satellite-terrestrial networks," in Proc. 3rd Int. Congr. on Ultra Modern Teleccommun. and Control Syst. and Workshops, Oct 2011, pp. 1–8.

[25] C. Morel, P. D. Arapoglou, M. Angelone, and A. Ginesi, "Link adaptation strategies for next generation satellite video broadcasting: A system approach," IEEE Transactions on Broadcasting, vol. 61, no. 4, pp. 603–614, Dec 2015.

[26] EU BATS Project, "Cost benefit analysis," Deliverable, April 2014. [Online]. Available: http://cordis.europa.eu/docs/projects/cnect/3/317533/080/deliverables/001-Ares20174277027317533BATSSubmissionofD52.pdf

[27] C. Niephaus, M. Kretschmer, and G. Ghinea, "Qos provisioning in converged satellite and terrestrial networks: A survey of the state-of-the-art," IEEE Commun. Surveys Tuts., vol. PP, no. 99, pp. 1–1, 2016.

[28] R. Maallawi, N. Agoulmine, B. Radier, and T. B. Meriem, "A comprehensive survey on offload techniques and management in wireless access and core networks," IEEE Communications Surveys Tutorials, vol. 17, no. 3, pp. 1582–1604, thirdquarter 2015.

[29] C. Sankaran, "Data offloading techniques in 3gpp rel-10 networks: A tutorial," Communications Magazine, IEEE, vol. 50, no. 6, pp. 46–53, June 2012.

[30] Y. He, M. Chen, B. Ge, and M. Guizani, "On wifi offloading in heterogeneous networks: Various incentives and trade-off strategies," IEEE Communications Surveys Tutorials, vol. 18, no. 4, pp. 2345–2385, Fourthquarter 2016.

[31] R. Trestian, K. Katrinis, and G. M. Muntean, "Ofload: An openflow-based dynamic load balancing strategy for datacenter networks," IEEE Transactions on Network and Service Management, vol. 14, no. 4, pp. 792–803, Dec 2017.

[32] S. Watts and O. Aliu, "5g resilient backhaul using integrated satellite networks," in Proc. 7th Adv. Satellite Multimedia Syst Conf. and 13th Signal Process. for Space Commun. Workshop, Sept 2014, pp. 114–119.

[33] W. Xia, Y. Wen, C. Foh, D. Niyato, and H. Xie, "A survey on software-defined networking," IEEE Commun. Surveys Tuts., vol. PP, no. 99, pp. 1–1, 2014.

[34] L. J. Chaves, I. C. Garcia, and E. R. M. Madeira, "Ofswitch13: Enhancing ns-3 with openflow 1.3 support," in Proceedings of the Workshop on ns-3. ACM, 2016, pp. 33–40.

[35] V. Hytönen, B. Herman, J. Puttonen, S. Rantanen, and J. Kurjenniemi, "Satellite network emulation with network simulator 3," in Ka and Broadband Communications, Navigation and Earth Observation Conference, 2014.

[36] J. Puttonen, S. Rantanen, F. Laakso, J. Kurjenniemi, B. Herman, O. Alanen, and K. Aho, "Simulator architecture and design definition," ESA, Tech. Rep., 2016.

[37] NGMN Alliance, "Next generation mobile networks radio access performance evaluation methodology," White Paper, 2008.

[38] [Online]. Available: http://www2.tkn.tu-berlin.de/research/trace/pics/FrameTrace/mp4/Verbose_StarWarsIV.dat

[39] E. Bocchi, L. De Cicco, and D. Rossi, "Measuring the quality of experience of web users," in Proceedings of the 2016 Workshop on QoE-based Analysis and Management of Data Communication Networks, ser. Internet-QoE '16. New York, NY, USA: ACM, 2016, pp. 37–42. [Online]. Available: http://doi.acm.org/2940136.2940138

[40] Estimating end-to-end performance in IP networks for data applications, ITU-T Recommendation G.1030, 2014.