Network slicing - enabled RAN management for 5G: Cross layer control based on SDN and SDR

Sokratis Barmpounakis*, Nikolaos Maroulis, Michael Papadakis, George Tsiatsios, Dimitrios Soukaras, Nancy Alonistioti

Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece

A R T I C L E   I N F O

Article history:
Received 31 July 2019
Revised 7 November 2019
Accepted 8 November 2019
Available online 9 November 2019

Keywords:
5G
Software defined networking
Network Slicing

A B S T R A C T

Having already entered the initial phase of the 5th generation networks (5G) era, the network ecosystems have already begun comprising a plethora of coexisting 3GPP and non-3GPP Radio Access Technologies (RATs). Deployment scenarios envision a multi-layer use of macro, micro and femto-cells, where multi-mode end devices, supporting different applications, are served by heterogeneous access technologies. At the same time, this heterogeneous environment results in an abstract pool of resources, moving away from the traditional cell concept and exposing the network administrator with the collective resources in time, frequency, and space. An imperative need is thus created, to obtain an overview of the network and radio conditions in order to respond in an optimal way. The concepts of Software Defined Networking (SDN) and Software Defined Radio (SDR), when applied in a coordinated and sophisticated manner, may prove of utmost importance towards addressing the aforementioned challenges. The concept of network slicing – that enables the creation of several virtual networks with diverse characteristics, on top of common infrastructure - further reinforces the flexibility and dynamicity of the network. To this end, we present and evaluate a Cross Layer Controller (CLC), which acts on top of well-established SDN and SDR technologies, and monitors in real-time the conditions in the radio environment, as well as the transport network traffic, in a unified way; based on this dynamic input, CLC enforces sophisticated, end-to-end resource allocation policies and pairing mechanisms between network slices from different network segments, highlighting thus, the feature of cross layer resource management and dynamic adaptation. The experimentation results demonstrate the considerable gains acquired from the proposed solution.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

The 5G wireless communication system aims to create a fully connected society, serving diverse business models, applications and services, linked with diverse traffic types and extreme requirements [1]. Furthermore, different networks and numerous User Equipment (UE) types will interact with diverse capabilities comprising Ultra Dense Network (UDN) deployments [2], thus making the overall future of wireless communications even more challenging. Meanwhile in the future, it is predicted that the traffic from the User Terminals will increase 1000 times from 4G to 5G [3]. Interference due to uncoordinated resource sharing techniques presents a key limiting factor in the design of such dense wireless networks, where resources are limited due to either the costs for licensed bands or the proliferation of hot spots in license-exempt bands.

Already, since the 4th Generation of networks, 3GPP has proposed the usage of LTE in unlicensed bands [4]. Towards 5G, it is expected that the future unlicensed LTE communications, i.e., LTE-U, will coexist in different types of indoor deployments with Wi-Fi deployments, creating critical interference challenges, which need to be handled effectively. 3GPP additionally, introduced Licensed-Assisted Access (LAA) [5] to provide operators and consumers with an additional mechanism capable of exploiting unlicensed spectrum for improved user experience, while coexisting with Wi-Fi or other radio access technologies (RATs) in the 5GHz unlicensed band. The backhaul infrastructure will inevitably need to operate in a seamless, RAT-agnostic manner, controlling the aggregated radio resources in a unified, flexible manner. As a result, an orchestrated control between the radio and the backhaul portion of the network resources is of utmost importance. This challenge requires the deployment of agile network controllers, orchestrating the util-

* Corresponding author.
E-mail addresses: sokbar@di.uoa.gr (S. Barmpounakis), nmaroulis@di.uoa.gr (N. Maroulis), mpapadakis@di.uoa.gr (M. Papadakis), george@di.uoa.gr (G. Tsiatsios), dsoukaras@di.uoa.gr (D. Soukaras), nancy@di.uoa.gr (N. Alonistioti).

https://doi.org/10.1016/j.comnet.2019.106987
1389-1286 © 2019 Elsevier B.V. All rights reserved.
lization of wireless and backhaul resources, to achieve efficiency in end-to-end (E2E) network performance, resource utilization, as well as energy consumption.

One of the cornerstones during the evolution of 4G networks towards the 5th generation of networks is the introduction of the network slicing paradigm, that enables the customization of segments of a common network infrastructure (i.e., “slices”) in order to meet the different requirements of the 5G use cases-, is one of the key enablers towards addressing numerous of the aforementioned challenges. Besides the high contributions of the relevant standardization bodies such as 3GPP, ETSI and IETF [6–8], Network Slicing is already one of the well-established enablers in the literature, and has already been proposed in numerous solutions [9,10].

However, although Network Slicing can “fine tune” specific portions of the network and optimize their performance according to specific business and service needs-, will not be always capable of addressing those needs in dynamic, heterogeneous wireless environments. To this end, besides the features brought by the network slicing paradigm, it is crucial for network management mechanisms to be able to monitor real-time network conditions for coping with abrupt, potentially unforeseen network traffic challenges and hazards, both on an inter- or intra- slice level. Network slicing should, thus, be applied in a flexible and dynamic manner, supported by real-time resource management schemes, especially in cases that network or radio resources are not abundant and the network must satisfy the on-going Service Level Agreements (SLAs). This work proposes such a real-time, cross layer resource control scheme in order to provide the required flexibility to the network slicing policies.

One of the recent approaches, -among which, also network slicing relies on-, is virtualization. Traditional virtualization work so far focused mainly on the network and computation level via enablers, such as the Network Function Virtualization (NFV) and SDN. With its origins in NFV, Radio Access Virtualization (vRAN) also offered further advantages, via centralizing part of the softwarized radio access point stacks into computing infrastructure in a cloud location [11]. Towards the same directions, Cloud RAN (C-RAN) and Multiple Access/Mobile Edge Computing (MEC) [12] virtualize RAN functionalities to a large extent, targeting to aggregate computing resources in a centralized location in the network; particularly, when combined in a paired manner, in a cloud-based approach, aggregated virtual computing resources near the network edge can have a direct impact on the network’s end-to-end latency for highly stringent network applications. Moreover, another aspect of such a cloud-based approach is that it enables a scalable solution, in particular making the capacity of the C-RAN dynamic.

In order to communicate with a wireless peer, logical/virtualization entities have to send and receive data via antenna. Programmable radio platforms, i.e. SDRs, can, thus, federate the existing approaches, as they are flexible enough to avoid the “limited spectrum” assumptions, to offer dynamic transmitter power adjustment, based on information communicated from the receivers, direction signal detection towards interference mitigation, cognitive radio techniques, spread spectrum and ultra-wide band techniques (i.e., very low energy level for short-range), etc.

In the research community, different implementations (e.g., srsLTE [13], OAI-LTE [14], open-LTE [15], GNU radio [16], WARP-WiFi [17], etc.) always ask exclusive access of programmable radio hardware (BladeRF [18], USRP [19], etc.). Hence, a challenge that emerges is how to share the same radio hardware/platform among those entities, while still maintaining the high efficiency and real-time operation of the system.

Unlike the computer/server or network switch/router models, -mainly considered as a computer system-, radio hardware/platforms are almost fully customized for specific standards, frequencies, bandwidths, etc. Therefore, radio resource control and real-time operation are challenging in the context of communication service virtualization. The proposed solution acts towards this direction, providing the capability to optimize in real-time the slicing approaches and policies between the network and the radio, where the most appropriate modulation, coding schemes and resource sharing mechanisms are selected and dynamic traffic steering policies can be applied.

The remainder of this article is organized as follows. The next section presents the state of the art. Then, we describe in detail the proposed framework in Section III. For the sake of performance evaluation, in the Section IV, we portray the experimental setup and discuss the obtained results. The article concludes in the final section with a summary recapping the main findings.

2. Related work

So far, several efforts have been proposed towards addressing the afore-discussed challenges via solutions that introduce the enablers of SDN and virtualization, either in the core of the network or the RAN. As discussed, SDN features flexibility and optimized resource allocation among heterogeneous services, enabling a plethora of business applications [20].

5G-EmPOWER [21] is a multi-access, OpenFlow protocol-enabled [22] Edge Computing OS, for SDN/NFV research and experimentation in mobile networks, which supports virtualization and is able to operate for heterogeneous radio access technologies. It provides high-level programming APIs, allowing prototyping of novel services and applications.

Another SDN-based architecture is proposed in [23], in which the authors propose a novel SDN-based, access backhaul network architecture, based on smart gateways between small cells and conventional backhaul gateways. Based on LTE protocol modifications, they target to reduce packet delays, via introducing flexible uplink packet transmission; at the same, SDN enables the management mechanism, where multiple operators, inter-operate via multiple smart small cell gateways and a multitude of small cells.

CrossFlow [24] introduces a cross-layer architecture platform, where heterogeneous networks comprising wired and wireless devices are managed in a uniform manner. The proposed testbed makes use of the AetherFlow [25], which enables wireless support in SDN, and adds additional API functions to the physical port abstraction.

In [26], the authors present a flexible way to optimize network resources sharing among wireless operators and nodes across the network by using gradient methodology. As a result, the traditional SDN resource management is extended into a distributed management module.

The study in [27] proposes a cross-layer architecture combining SDR and SDN principles. The architecture comprises two main layers: the SDR and the SDN layers. The key element of this mechanism is the cooperation between SDN and SDR such that SDR can have access to monitoring results of frequency spectrum and SDN can use the frequency spectrum conditions provided by SDR, when policy has changed. Nevertheless, the authors evaluate the proposed mechanism in a simplified, MATLAB-based simulation, generating positive results with regard to bandwidth utilization rate.

In [28–30], the authors propose different architectures of software-defined heterogeneous networks, based on the integration of NFV, SDN and SDR technologies. SDR hardware is used for providing programmable access infrastructure, OpenFlow-based SDN is exploited towards flexible access control and bandwidth management, while NFV techniques are also deployed in order to accomplish flexible service chaining. In [28], the authors propose a concept of functional blocks of control and monitoring of the QoS, that is realized as a part of a cloud infrastructure, namely the Cloud QoS Monitoring Function (CQMF); however, no detailed
evaluation is presented in this work. The work in [29] applies virtualization on currently deployed physical SDN networks. The proposed study suggests integrating both wireless and optical domains; however, the authors illustrate the advantages of the proposed concept via a limited, simulated model, comprising only a wired network of switches and end-hosts, using iPerf tool between them for monitoring the achieved data rate, missing to demonstrate potential gains in the wireless portion of the network. The work in [30] describes a conceptual architecture exploiting SDN, SDR, and NFV technologies, however, no evaluation is present to validate the proposed concept.

Instead of concentrating data and computation in a small number of large clouds, many edge systems are envisioned to be deployed in the vicinity of the end users or where computing and intelligent networking can best meet user needs; to this end, the authors in [31] propose an architecture called MUREN (Multi-Radio Edge Node) for managing traffic in future mobile edge networks. The proposed solution is based on the Mobile/Multiple Access Edge Cloud (MEC) architecture and SDN, both jointly interacting with SDRs. The prototype is implemented on a real testbed using USRPs [19] and GNU Radio [16], while the SDN/NFV part is based on Open vSwitch [32].

Last but not least, the H2020 European Project’s ORCA [33] primary objective was to enable E2E real networking experiments combining SDR with SDN and Dynamic Spectrum Sharing (DSS). Towards bridging SDN and SDR technologies, the project aims the creation of multiple virtual networks that operate on the same infrastructure but meet the most diverse and stringent application requirements.

Some excellent related work has been presented so far in this section, which covers to a considerable extent several of the discussed challenges, via introducing SDN, SDR and virtualization features into the network architecture. Some of the proposed solutions prioritize the SDN aspect of their solutions, targeting to showcase how centralized control of the resources may prove of utmost importance towards optimal resource management; others primarily focus on SDR implementations, while some conceptually present an integration of SDN and SDR, lacking however realistic evaluation scenarios, as the majority of them rely on simulations. Additionally, the reviewed solutions do not present details in their approach on the policies’ enforcement methodology, particularly for policies, rules and “events” orienting from different segments of the network. Last but not least, most of the proposed solutions have been designed based on specific software (i.e. SDN/SDR tools and protocols) lacking flexibility in terms of real network deployments, with already deployed SDN and SDR technologies.

The proposed work attempts to provide a holistic methodology, comprising: a flexible framework, capable of integrating any 3rd party controller (either SDN controller, SDR controller software, etc.), highlighting thus, the importance of flexibility in such control frameworks, -something that is missing in the existing work to the best of our knowledge-; secondly, a detailed description of the methodology and policy enforcement algorithm is provided in detail, in contrast to the majority of the afore-presented solutions, which mainly highlight the functional and conceptual aspect of their solutions; last but not least, to our knowledge, no previous work has been evaluated to such an extent in a real world 5G testbed, i.e. NITOS Wireless testbed [34], as part of the Fed4FIRE+ federation [35].

3. Cross layer control based on SDN and SDR technologies

In this section, we present CLC, a Cross Layer Controller that takes into account the RAN measurements, as well as and network-specific information in a unified way and in real-time, and applies policies via flexible, dedicated interfaces with 3rd party “sub-controllers” towards the different portions of the network.

3.1. The resource pool

The radio resources are considered as an abstract resource pool that moves away from the cell concept and exposes the network administrator with the collective available network resources in terms of time, frequency and power. Fig. 1 illustrates the dimensionality of those resources as perceived by CLC, and -as a result- the flexibility that is created towards the actions’ enforcement, that will be described in Section III. The power dimension corresponds to the dynamic transmitter power adjustment capability, which is used in SDR technologies towards minimizing interference to nearby wireless elements, the time dimension refers to the temporal aspect of the resource allocation and the duration of the allocated resource elements per UE, while the frequency corresponds to the dynamic, multi-channel/bandwidth capabilities of the transmission.

The minimum allocation unit in the resource pool is the resource block and is identified by a frequency band, a time interval and the radio access point (e.g., Wi-Fi AP, eNB, LTE femtocell, etc.) that provides it. The radio resource pool is managed by the SDR Controller, an entity capable of selecting the appropriate radio configuration, depending on the available radio context and service requirements. Thus, managing radio resources in a unified manner improves not only radio resource management, but resource management as a whole, as backhaul network links are also mapped to respective radio resources in the access network in an end-to-end manner, formulating thus, end-to-end network slices; in a similar manner to the radio resource pool, the network resources –comprising all the transport network links under the SDN controller’s overview- are also managed in a unified manner in the control plane.

Following the network slicing paradigm [36], slices from the different network segments and the respective resource pool, i.e., from Core Network (CN) to Radio Access Network (RAN) to radio link –mapped with each other comprising an end-to-end slice-. An end-to-end network slice includes one CN slice, one or a few RAN slices and one or a few radio slices. CN and RAN slices need to be set up in order to provide certain CN services and procedures and RAN capabilities, such as low power wide range access, low latency high reliability access, specific access base stations, etc.; radio slices refer to tailored radio resources with adaptive scheduling and HARQ framework, flexible duplexing modes, numerologies, etc. (e.g., tailored for MTC, V2X slices, Narrow-Band IoT, etc.).

CLC focuses on the RAN and radio slices adaptation and enforces E2E resource management policies in a both intra-slice (i.e., slice characteristics adaptation), as well as the inter-slice manner (i.e., slice pairing between RAN and radio slices); those policies are then applied exploiting the SDN and SDR controllers’ capabili-
ties via the respective interfaces. Such capabilities comprise adaptive traffic steering rules of network flows, the power control of specific Wireless Termination Points (WTPs) (LTE eNBs, Wi-Fi APs, etc.), spectrum allocations policies, as well as the resource blocks’ scheduling schemes.

3.2. CLC architecture and design principles

As discussed earlier, one of the key architectural approaches in CLC is to provide the necessary abstraction in order to enable any 3rd party controller or radio protocol to be integrated in the proposed architecture, increasing thus, its flexibility and potential for real network deployments, with already-deployed controllers, either for the SDN part, or for the respective programmable radio devices. With regard to the data plane, CLC design follows the principle of adding additional API functions for accessing the various programmable radio components [37]. This categorizes the interfaces into a) the configuration interface (allowing the controller to configure the properties of the programmable radio device such as transmission power, frequency, gain, etc.), b) the capabilities interface (allowing the controller to obtain capabilities of the radio such as number of channels, supported bandwidth, etc.), and c) the monitoring interface (allowing the controller to aggregate information and statistics from the radio).

Fig. 2 illustrates an instantiated CLC example deployment, operating on top of a) Ryu framework, which uses OpenFlow protocol [38] for the SDN part, i.e., the management of the network flows, b) the EmPOWER [21] controller for the Wireless SDN part of the network, as well as c) the Open Air Interface (OAI) platform [14] for controlling the programmable radio (SDR) part, which is based on Ettus Research USRP devices [19].

The core module of CLC is the CLC-Logic module, where all the available network and radio information is aggregated and all the decisions with regard to the policies and rules that will be applied in the network are taken. A conceptual representation of the aforementioned mechanism is illustrated below (Fig. 3):

Table 1 demonstrates the radio-related metrics, network information and other types of context information, which are monitored and forwarded to the CLC for further processing and decision making.

3.3. The policy-rule-action model

As already discussed, CLC is constantly receiving context information (see Table 1) from the various network elements and events are triggered. CLC then follows a Policy – Rule/Event – Action mechanism that is described in detail below.

Semantically, a Policy is linked to a number of Rules, while a Rule is mapped to a number of Events and their respective Actions (Fig. 4).

A Policy is a set of Rules that target a specific set of KPIs and a set of network services within a network slice. Usually, the Rules that comprise a specific Policy, act in a supplementary and federated way in order to accomplish the specific policy, defined by the network administrator. Policies are manually defined by a network administrator according to specific use cases requirements’ that need to be addressed, existing SLAs, etc. and is applied for specific network slices.

There may be numerous defined Policies, -even conflicting, depending each time period on the specific slice/service KPIs and respective SLAs-, but only a number of them is activated at a time period by the network administrator, towards federating/aligned objectives. At some point, a Policy may be deactivated, and a different one –with previously conflicting objectives- may be activated.
As a result, at any given moment in time, conflicting/antagonistic actions are avoided.

A Rule is a connection of Events, which are evaluated based on the input parameters, and a number of respective actions to be taken if these events are triggered. When a Policy is activated by a network administrator (i.e., to be applied in the network operation), the Policy’s Rules are also activated, along with their respective Events.

An activated Event means that it is in a iterative state of evaluating the received context and can be triggered at any time. When all Events of a Rule are triggered, then the Actions of the respective Rule are enforced.

The command list of Actions is a predefined custom set of script-based commands, based on the available platform and controllers’ capabilities (RYU API, OAI radio API, etc.).

The CLC algorithm and actions’ enforcement methodology description follow in detail:

Let us denote:

\( P: \) the set of all available defined policies,
\( P,R: \) the set of rules, linked to Policy \( i \)
\( P,R,E: \) the set of events, linked to Rule \( j \) of Policy \( i \)
\( P,R,A: \) the set of actions linked to Rule \( j \) of Policy \( i \)

\( p_i^j = 1: \) binary value equal to 1 if policy \( i \) has been activated,
\( p_i^j,r_j^k = 1: \) binary value equal to 1 if jth rule of the respective activated policy \( i \) was activated,
\( p_i^j,r_j^k,e_m: \) binary value equal to 1 if mth event of the jth rule of the ith policy is triggered,
\( p_i^j,a_n: \) nth action of the jth rule of the ith policy

The algorithm is defined in pseudocode as follows:

1. \( p_i^j, = 1 \) (Policy \( i \) activated by the network administrator)
2. \( \text{VisP IF } (p_i^j, ==1) \text{ THEN } p_i^j,a_n^j == 1 \quad \forall \text{inPLR} \)
3. \( \text{while } p_i^j, ==1 \)

/* this step refers to the iterative loop of monitoring the events triggering, as long as the respective policy is activated */

\( \text{VisE, VisPLR IF } (p_i^j,r_j^k, ==1) \text{ THEN } \\
\text{a. IF } (p_i^j,r_j^k,e_m, ==1) \text{ THEN } \\
\text{DO } p_i^j,r_j^k,a_n \quad \text{execute all actions of specific triggered rule} */

In order to provide an example of the CLC Policy-Rule-Action model, Table 2 illustrates some simple cross layer Rules and Actions, both on the network and the radio side:

### 3.4. Action enforcement

Fig. 5 illustrates the mechanism of the policies’ and rules’ formulation and the respective mapping to specific radio slices. It must be noted, that although the actual rules and action enforcement to the network devices is applied per network segment, effectively the optimization is applied in end-to-end manner, as the Rules and respective Actions of a single Policy concern all the different network segments in correlated manner.

With regard to the Actions’ enforcement that satisfy each policy, the framework provides three different approaches, i.e., a) the manual operation mode, b) the automatic operation mode and c) a semi-automatic (or “hybrid” mode), in which both manual and automatic action enforcement may take place, and specific actions may be paused. As already discussed, the formulation of policies and rules is manually defined by the network administrator, who is aware of the existing SLAs, the requirements of the domain-specific network slices and vertical services KPIs, their prioritization, as well the time periods that each one of the aforementioned services must be (de)activated. However, the actual enforcement of the Actions –as described in Section III-C-, when policies, rules and events are triggered, may be applied either automatically (using script-based tools like Ansible [40]), according to well-established Automation and Orchestration frameworks’ procedures –or manually–. As it will be shown in the Evaluation section, for the purpose of the specific evaluation, and in order to monitor in detail the various actions’ effects as the experiment progresses, we used the semi-automatic approach.

During the Policies and Rules definition, the network administrator, uses a customized user interface (Fig. 6–Fig. 8) in order to formulate the desired policies, or edit specific characteristic of existing radio slices (e.g., allocated bandwidth, deployed VNFs, etc.). Via the CLC Graphical User Interface (GUI) or by using command line tools (both well-known network tools such as ping applications, as well as special API provided by the framework controllers, such as OpenFlow API, etc.) the administrator monitors in real-time the condition and efficiency of the network links, the status of the radio conditions, the load of the base stations/access points, etc. and (de)activates Policies and Rules according to the desired strategy.
Table 2
Example cross layer policies and commands.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Description</th>
<th>Rule</th>
<th>Events</th>
<th>Action #</th>
<th>Action Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Offload macro base stations</td>
<td>1.1</td>
<td>IF (connected_to_macro-1) &amp; &amp; (macrocell_load_condition_C) &amp; &amp; (femtocell_load_condition_C) &amp; &amp; (backhaul_datapath_C) &amp; &amp; (available_femtocell) &amp; &amp; (L2_switch_command_COM)</td>
<td>1.1.1</td>
<td>UPDATE (12_switch_command_COM_i)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2</td>
<td>IF (connected_to_macro-1) &amp; &amp; (macrocell_C) &amp; &amp; (femtocell_C) &amp; &amp; (L2_switch_command_COM)</td>
<td>1.2.1</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.1</td>
<td>IF (link_S_load_C) &amp; &amp; (RSRQ_Ui) &lt; R, &amp; &amp; (num_neighbour_BS) &gt; N &amp; &amp; (L2_switch_command_COM)</td>
<td>2.1.1</td>
<td>UPDATE (12_switch_command_COM_i)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.2</td>
<td>...</td>
<td>2.2.1</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td>...</td>
<td>2.3.1</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>Congested backhaul link</td>
<td></td>
<td>...</td>
<td>3.1.1</td>
<td>INCREASE BS, TX_POWER = 10%</td>
</tr>
<tr>
<td>3</td>
<td>High Interference</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Fig. 6. Main CLC GUI screen illustrating the set-up topology.

3.5. The “Stadium” use case

In order to define and illustrate CLC’s design and functional requirements, we designed a conceptual use case (Fig. 9) that demonstrates the capabilities of CLC and evaluates its performance. The use case illustrates portions of a realistic network topology of an urban environment comprising diverse RATs in different areas of the city, all interconnected via a common backhaul based on programmable switching and routing infrastructure.

CLC is capable of monitoring in real-time the overall network’s conditions, receive real-time information from the programmable base stations (BSs) (macro/pico/femto cells, etc.) regarding their loads, available bandwidth, number of connected UEs, etc. In this scenario, we assume several potential network issues and events that may hinder the performance of the network.

The main event that influences that the overall network performance of the area is a football match that takes place at a specific area of the city (Fig. 9, top left corner); the “Stadium” use case is considered as one of the well-established 5G use cases as described by one of the pioneering 5G project, i.e., METIS 2020 [39]. The stadium is equipped with several pico and femto cells, as well as 802.11 APs, creating an Ultra Dense environment, when a high number of clients is also connected to the BSs inside and outside the stadium. Moreover, several residents are living nearby the stadium and they are often connected to these BSs during weekdays (where there is no event).
We focus on the consequences that a Sunday match has on the overall network operation, along with the potential strategy that could follow using CLC framework. A high number of UEs (on the order of tens of thousands) starts approaching and requesting access via the stadium’s base stations, already a few hours before the beginning of the event. As a result, the femto cells of the stadium and the pico/macro BSs outside the stadium are beginning to accommodate more and more UEs, which may request high bandwidth applications (e.g., HD video streaming). CLC monitors this gradual connection increase a) on the backhaul link that connects the stadium area’s BSs and APs to the provider infrastructure and b) on the residents that make use of these BSs normally during weekdays.

As a result, a new policy prioritizing the stadium UEs for the specific BSs starts being applied using CLC. Handover commands are forwarded to the residents’ UEs (Fig. 9 – yellow lines), switching their active sessions either to their home APs or nearby 4G/5G BSs, which are not utilized by the stadium UEs, in order to prevent any possible degradation of their Quality of Service (QoS). At the same time, there is a timely update of the backhaul data paths (green lines) of the respective residents’ APs, in order to detour the soon overloaded- stadium’s backhaul link (purple line).

Finally, the use case involves a scenario where a potential emergency call is required to take place from the stadium area in case of an urgent event (e.g., a fire breaks out) and the fire brigade needs to be notified immediately. In such a case, a dedicated network slice is utilized. Therefore, the emergency call data flows are localized and specifically forwarded via the alternative data (red line) path towards the fire brigade location (bottom centre of the figure), in contrary to the rest of the flows coming from the stadium BSs and APs.

The selected scenario is used to describe how information, received from different parts and layers of the network (radio access inside the stadium, backhaul network links, emergency/high-priority service requests, etc.) can be further processed, towards E2E policy enforcement, and QoS optimization for different types of UEs and types of services. Based on this use case, two sub-scenarios -formulated in order to demonstrate parts of the afore-described CLC’s operation- are presented in the next section.

4. Evaluation

In order to assess the performance and validity of the proposed framework, extensive experimentation took place in one of the Fed4FIRE+ testbeds [35], i.e. NITOS Wireless testbed [37] (Fig. 10). During the evaluation, CLC operated in the semi-automatic mode, i.e. the approach of automatic real-time monitoring and automatic actions’ triggering was followed; nevertheless, in order to monitor the efficiency and outcome of each one of triggered actions, we paused the serially executed actions and investigated each one of the intermediate steps, on a standalone manner. A detailed description of the two main scenarios that were evaluated follows in the next subsection.

4.1. Scenarios’ description

4.1.1. Scenario 1

The main idea of the first experimentation scenario is to improve specific network KPIs, i.e., the downlink and uplink throughput, the downlink and uplink jitter, as well as the experienced interference, via processing and correlating context information received from both the radio and the network layer (link utilization information) of the network. CLC is used in order to timely identify potential congested network links and overloaded radio channels and consequently take the appropriate actions, in order to improve the aforementioned KPIs. The actions taken, include -but are not limited to- handover commands towards load balancing the available radio channels/wireless access points, reconfiguration of radio transmission parameters (such as Tx/Rx power, etc.) as well as updating backhaul data paths and re-routing network flows, in order to mitigate the network congestion, and as a result- ultimately enhance the system’s E2E performance for the particular scenario’s objectives. Especially, traffic rerouting and slice isolation for the emergency slice (Fig. 9, red line), based on coordinated actions on multiple portions of the network, is one of the key objectives. Regarding the topology of this scenario, two programmable WTPs (Wi-Fi APs) were set in NITOS testbed, along with five wireless clients, each connected to one WTP. Furthermore, three OpenFlow-enabled switches were controlled using the RYU controller API and the EmPOWER Controller. Fig. 11 illustrates the afore-described topology and main scenario steps.

4.1.2. Scenario 2

The second scenario (Fig. 12) focuses on the coexistence of LTE and 802.11 technologies –operating in overlapping frequency bands-, aiming to simulate a dense, heterogeneous wireless environment. The main rationale behind this scenario is to demonstrate the feasibility of having a single point of control between heterogeneous coexisting Radio Access Technologies (RATs), where all the information is aggregated and the policies are taken in a unified and holistic manner, considering the resources of both the LTE as well as the Wi-Fi networks as part of the overall abstract resources pool.

Similarly with the first scenario, the experiment was also carried out on NITOS testbed, where we reserved programmable radio nodes (based on the USRP B210), hosting the OAI 5G Controller.

![Fig. 10. NITOS Indoor RF Isolated testbed.](image-url)

![Fig. 11. Scenario 1.](image-url)
software, nodes connected with HUAWEI E3372 LTE Dongles and one generic node hosting the OAI Evolved Packet Core (EPC).

4.2. Results

4.2.1. Scenario 1

As already discussed, in the first series of the experiments, the main goal is to correlate radio congestion information (high channel utilization, interference, etc.) with information from the network flows routed via the switches in the backhaul and then apply the required actions, in order to ameliorate specific KPIs, both in the radio as well as the backhaul. The radio devices in this scenario are Wi-Fi APs.

A high number of clients, downloading at max available throughput using an SFTP service, connect to AP1, creating a high radio channel utilization. For performance reasons, we used the capability of SFTP for file transfer clients to be able to send multiple requests without waiting for responses. Using the Ryu controller, we identify the respective congested links in the backhaul. Fig. 13 illustrates the Busy ness metric, an EmPOWER-related KPI, that correlates the overall bandwidth with the measure channel interference. Fig. 13 shows the initial status of the network: a high number of clients are connected to one of the two APs, that operate in the same frequency channel. Our test UE/Wi-Fi client (in red) – initially attached to the congested AP – experiences very high busy ness even when idle, due to the highly congested environment. A handover is realized to the 2nd AP, in order to ameliorate the test client’s KPIs. However, minor changes are observed. The overwhelmed test client’s status is also proven by the bandwidth measurements in the backhaul. Both Wi-Fi APs are initially routed via the same switch, which results in bottleneck effects on the link layer as well.

Fig. 14 illustrates the outcome of the first action, which involves creating a radio slice, isolated and interference-free from the congested radio channel, and its adjacent ones. In this phase, we switch the 2nd AP’s operating frequency band, on the other edge of the Wi-Fi spectrum, clearing out all radio interference. This time, the test client’s handover to the sliced, 2nd AP has a clearly positive effect. However, the problem remains in the backhaul. Although the new channel is interference-free, both APs’ outgoing and incoming traffic is routed via the same switch, sharing a common backhaul.

Fig. 15 illustrates the effect of the 2nd action taken. Using the Ryu/OpenFlow CLC’s interface we create a second slice, which maps all the flows coming from going to the 2nd (test client’s) APs to be forwarded by a new network route. The bottleneck removal shows immediate positive effects both the radio part’s KPIs, as well as the Layer 2 and 3 results, which are ameliorated considerably (measured by Ryu/OpenFlow traffic monitoring tool). This proves that although the SDR is crucial for interference issues related directly to the radio segment, actions in Layer 2 and 3 (via targeted and selective rerouting of specific network flows in congested backhauls of specific RAN components) are also vital for an end-to-end optimization.

4.2.2. Scenario 2

The main objective of the 2nd scenario was the improvement of the network’s status in a coexisting unlicensed LTE-Wi-Fi environment, with a gradually increasing number of wireless hosts being connected to each RAT respectively, correlating the KPIs measured from both Access Technologies and applying the respective policies.
Regarding the traffic type that was used, similarly with Scenario 1, the hosts are downloading at max available throughput using an SFTP service. The two RATs operate in overlapping frequencies, and as a result, interference and band occupation issues have to be addressed. In the beginning of the experiment, only one testing UE is connected, while gradually, during the experiment, 30 hosts get connected -per RAT-, in a limited room environment of a room with UDN characteristics. CLC’s OAI module is used for sensing/actuation in the LTE part of the network, while the EmPOWER module for the Wi-Fi. CLC performs eNB frequency shifting, coordinated power control operations and handover commands. The following figures (Fig. 16 and Fig. 17) illustrate how the overall system’s average throughput and jitter (latency variation) results for a specific testing host respectively vary in a step-by-step manner, after each CLC applied action.

The experiment is described step-by-step (as also illustrated) in the above figures:

- **Step 1 (S1):** initial state – single eNB/UE connection – no interference is identified

- **Step 2 (S2):** We deploy 30 UEs that connect to the available LTE cells. Due to the coexisting UEs the interference is increased and thus the LTE UEs overall performance is considerably low, i.e., decreases by 10–40%

- **Step 3 (S3):** 30 Wi-Fi nodes are also deployed in the same area. The Wi-Fi is operating in an overlapping channel with the unlicensed LTE band so performance on the overall system deteriorates further

- **Step 4 (S4):** CLC sends a handover command to the testing UE, moving it to an interference-free and less-loaded LTE cell – throughput and delay are improved

- **Step 5 (S5):** Coordinated power control is applied for the LTE cells in order to boost the testing UE performance by increasing the Signal-to-Noise Ratio (SNR) and respectively decrease the Bit Error Rate (BER) metric – however, heavy deterioration is monitored on the neighbor UEs that must be addressed

- **Step 6 (S6):** Wi-Fi APs channels, as well as the LTE cells’ frequencies are switched to interference-free bands after identifying the frequency overlap. Performance is highly improved.

The provided, step-by-step experimentation proves the validity of this real-time, monitoring and action-generating approach, showcasing considerable improvements in real-life scenarios, such as coexisting heterogeneous RATs.

Overall, in this section, the two scenarios we carried out, lead to the following conclusions:

- The management of dense wireless deployments towards 5G requires a holistic view of the available resources, comprising backhaul infrastructure (switches, links, etc.), RAN infrastructure (eNBs, LTE femto cells, Wi-Fi APs, etc.), as well as spectrum conditions

- The experimentation that was carried out in NITOS proves that dynamic radio resource management using network slicing mechanisms, and exploiting wireless SDN and SDR approaches, has a direct effect on the measured performance KPIs

- When combined with coordinated actions related to the backhaul network (e.g., dynamic flow management using OpenFlow switches) a higher enhancement of these KPIs is reported.

5. Discussion and conclusions

5G-network management networks should simultaneously optimize coherence between the control plane and data plane, while minimizing their coupling and employing principles for dynamic and flexible operation. The convergence of the SDN’s programmable control plane nature with the SDR’s programmable data plane into a novel architecture will thus enhance and complement both scopes; this is the core concept of this work, i.e., the SDN-SDR coordination and their cross layer control architecture. Some example applications of this architecture comprise – but are not limited to – physical layer adaptation (frequency hoping, transmission power control, etc.), Quality of Service (QoS) provisioning to allow specialization of QoS policies through medium access control, adaptive routing to allow a cross-layer controller with a global view of the network to switch between routing protocols, cross-layer dynamic network slicing to enable optimized delivery of services, etc.

To this end, in this paper, a Cross Layer Control framework was introduced, which exploits SDN and SDR techniques in order to optimize the network performance in dense radio environments comprising heterogeneous radio access technologies. A detailed overview of the framework was presented, its architectural principles, as well as the main CLC concepts, such as the Policies, Rules and Actions enforcement methodology. Furthermore, a detailed use case was provided, defining in detail the functional and
technical requirements of the proposed framework and highlighting the importance of a holistic view of the network's underlying infrastructure. The extensive experimentation - using real programmable hardware - in an advanced wireless testbed proved the validity of the mechanism.

One of the primary targets for the continuation of the specific work is to further elaborate on the network slicing mechanism of the CLC framework and include Virtual Network Function placement and chaining policies. Secondly, we will elaborate on the how the proposed cross layer architecture influences the end-to-end latency, session set up and data flow: CLC builds upon the existing SDN architecture, relying on well-established SDN controllers and SDR technologies for its instantiation, no additional overhead is posed to the end-to-end operation delay. The CLC control - according to the standardization - operates on the application layer of the SDN architecture, similarly with load balancing, network management, policy enforcement apps, etc. Nevertheless, we will investigate in higher detail to what extent the proposed architecture influences the existing procedures.

Another important aspect that will be investigated in a follow-up work relates to the metric of fairness during the resource allocation rules formulation and the respective action. In general, increasing the overall system throughput while at the same time maintaining fairness is a tradeoff problem between the two metrics. In the proposed work, the objective is not to maximize throughput, as a generic resource allocation problem formulation; the primary objective, taking into account the specific network service KPIs, is to satisfy all KPIs (uplink/downlink throughput, delay, packet loss, etc. - each time depending on the service-) via ensuring that the policies and rules are satisfied. Due to the actions that are thus applied, the different network services' and slices' performance constantly converges towards the target KPIs, as also indicated by the evaluation. Nevertheless, investigating the fairness metric between different network slices, or different network services is of high interest and will be worth investigating in one of the follow-up works.

Last but not least, we plan to apply an advanced Artificial Intelligence-based model, exploiting Deep Reinforcement Learning techniques in order to further improve the framework's reconfiguration capabilities and carry out a new, extensive round of experiments in NITOS testbed, in order to compare its performance, with the results presented in the current work.

Declaration of Competing Interest

None.

Funding

This research was funded via Fed4FIRE+ project, under grant agreement No 732638 from the Horizon 2020 Research and Innovation Programme, which is co-funded by the European Commission and the Swiss State Secretariat for Education, Research and Innovation.

Supplementary materials


References

[34] NITLab, Network implementation testbed laboratory, https://nittahlab.inf.at/g/h/ NITLab/nitlab, [accessed 07/2019].
Dr. Sokratis N. Barmpounakis is a Postdoctoral Research Associate in SCAN group, Dept. of Informatics and Telecommunications, in the National and Kapodistrian University of Athens (NKUA). He holds a Ph.D. in “Context-based Resource Management and Slicing for SDN-enabled 5G Smart, Connected Environments”, since May 2018. Dr. Barmpounakis obtained his Engineering Diploma in Electrical and Computer Engineering, from the National Technical University of Athens (NTUA) in 2010. After completing his studies, he moved to Geneva, Switzerland. At University of Geneva, he worked for 2 years as a Researcher/Developer in various projects, Swiss and European, as a member of the Qol group. Since March 2013, he is a researcher in the SCAN group, actively involved in numerous European Projects and industrial contracts, having also undertaken managerial roles, coordinating the University of Athens team. Dr. Barmpounakis is a co-inventor of 2 patents, approved and filed in the World Intellectual Property Organisation. His main fields of interest are: 5G, LTE, Software Defined Networking, Context-Aware Resource and Mobility Management and Collaborative Data Sharing mechanisms. He is also a member of the Technical Chamber of Greece.

Nikolaos Maroulis obtained his four-year Bachelor of Science (B.Sc.) from the Department of Informatics and Telecommunications at the National and Kapodistrian University of Athens (UoA), with specialization in the Telecommunications and Networks field. He is currently an M.Sc. student at the same department, in the field of Data Science, while working as a research associate at the Software Centric and Autonomic Networking (SCAN) Lab of UoA. His main fields of interest are Software-Defined Networking (SDN), especially in the Wireless Domain (SD-WLAN), web technologies, distributed systems, Internet of Things applications based on SDWLAN, Software Defined Radio (SDR), 5G.

Michael Papadakis graduated from the the Department of Informatics and Telecommunications, NKUA, and is currently completing his M.Sc. at the University of Pisa, Dept. of Digital Systems, Object Oriented Architecture and Services. He is collaborator with SCAN group, University of Athens and with TERRA SPATIUM, Geo-information and Space Products & Services S.A. His field of studies include Cloud Computing, SDN, SDR, 5G communications, as well as conceptual models, system analysis, decentralized architectures and smart maritime applications. He is currently researching on computer vision and deep neural networks on real time communications.

George Tsilias obtained his four-year Bachelor of Science (B.Sc.) from National and Kapodistrian University of Athens (NKUA), in the Department of Informatics and Telecommunications, with a specialization in Telecommunications and Networks. He is currently a student at NCSR Demokritos, in the field of Data Science and works as a Researcher at the Software Centric & Autonomic Networking (SCAN) Lab of UoA. His main fields of interest are AI, Data Analytics, Software-Defined Networking (SDN), Software-Defined Wireless Local Area Networking (SDWLAN), Future networks and 5G.

Dimitrios Soukaras obtained his Bachelor of Science (B.Sc.) from the Department of Informatics and Telecommunications of University of Patoponese (UOp). He is currently an M.Sc. student at Department of Informatics and Telecommunications of National and Kapodistrian University of Athens (UoA). He served as a software engineer during his bachelor internship in ARIES team of INRIA Paris in the field of Internet of Things and sensor networks. He serves as a research fellow in the Software Centric & Autonomic Networking (SCAN) team of the Department of Informatics and Telecommunications of the National and Kapodistrian University of Athens (UoA) under the supervision of Prof. Nancy Alonistioti since 2013. He has been involved in several EU projects including: Flipspace, FRACtALS and SPECIFI. He serves as a system administrator of SCAN team’s in-house cloud infrastructure. His main research interests are Internet of Things, Software-Defined Networking, cloud management and LTE in 5G.

Assoc. Prof. Nancy Alonistioti is an Associate Professor in the Department of Informatics and Telecommunications, NKUA. She has a B.Sc. degree and a PhD degree in Informatics and Telecommunications (Dept. of Informatics and Telecommunications, NKUA). She has over 20 years of experience in numerous national and European projects, including project/technical management experience. She is currently leading the SCAN group activities in the Dept. of Informatics in the National, NKUA. She has served as member of the Future Internet Assembly Steering Committee. She has also undertaken management activities as a PMT member and WP Leader of the FP6 IST- E2R and FP7 E3 projects and the PC of FP7 ICT Self-NET and FP7 ICT CONSERN projects. She is member of the ETSI Experts group and the Greek standardization group ELOT (5G, smart city autonomic communications). She has over 100 publications in the area of mobile networks, Future Internet/NGI, cognitive management, autonomic communications and reconfigurable mobile systems. She is co-author of 4 WO Patents and has more than 2500 citation in her work.