RAN slicing for multi-tenancy support in a WLAN scenario

Katerina Koutlia, Anna Umbert, Sergio Garcia, Ferran Casadevall
Dept. of Signal and Theory Communications (TSC)
Universitat Politécnica de Catalunya (UPC)
Barcelona, Spain
{katkoutlia, ferranc, annau, sergio.garcia}@tsc.upc.edu

***Abstract*—Radio Access Network (RAN) slicing is a key technology, based on Software Defined Networks (SDN) and Network Function Virtualization (NFV), which aims at providing a more efficient utilization of the available network resources and the reduction of the operational costs. On that respect, in this demo a Wireless LAN hypervisor is presented that is based on a time variant scheduling mechanism and that is able to follow the dynamicity of the traffic variations seen by the different tenants in the network Access Points (APs). The work builds upon the 5G-EmPOWER tool kit, which is provided with SDN and NFV capabilities. During this demo it will be shown that the proposed hypervisor is able to dynamically assign, in every AP of the network, the appropriate resources per tenant according to their traffic requirements.**

Keywords—SDN-NFV; Virtualization; Hypervisor; RAN slicing; 5G-EmPOWER testbed;

I. Introduction

Current network challenges consist on their flexible and efficient management, while keeping the capital (CAPEX) and operating (OPEX) expenditures low. In order to meet up with these needs, virtualization played a crucial role since it decouples software applications from the underlying hardware [1]. To this end, SDN and NFV are considered as key technologies that introduce flexibility in the network and resource management [2]. With SDN and NFV the available resources can be shared dynamically and the RAN can be sliced according to the specific requests. This can be particularly beneficial in order to address the challenges imposed by the traffic demand in wireless networks. Each Virtual Network Operators (VNO) or tenant, may have different requests in a particular AP over time depending on the traffic variations on their network. RAN slicing allows to deploy and operate multiple logical networks over a common physical network infrastructure in a cost effective way, allowing each tenant to have its own logically isolated slice of resources with its own desired set of services and the complete control of them

In order to guarantee resource isolation between r tenants, a hypervisor must be introduced at the virtualization layer. However, in most of the cases, a fix share between the assigned resources to each tenant is assumed. This contribution employs a new strategy that exploits the concept of virtualization to provide a flexible resource allocation per AP. It will be demonstrated that this new hypervisor targets to maintain the SLA in a long term perspective and considering all the APs of the network, while satisfying traffic demand fluctuations in the short term in the individual APs. The demo builds upon an experimental platform that provides SDN and NFV capabilities, known as 5G-EmPOWER [3].

II. Demonstrator Overview

The implementation of the proposed architecture is based on the 5G-EmPOWER test-bed. The core of the test-bed is composed by the EmPOWER Controller and multiple Agents running on each AP (see Figure 1). From now on, by using the terminology of the EmPOWER, each AP will be referred to as Wireless Termination Point (WTP). The Controller allows multiple virtual networks (i.e. slices) to be instantiated on top of the same physical infrastructure. Moreover, it ensures performance isolation among the different slices. Network Apps of a given virtual network are implemented on top of the controller in their own slice of resources.

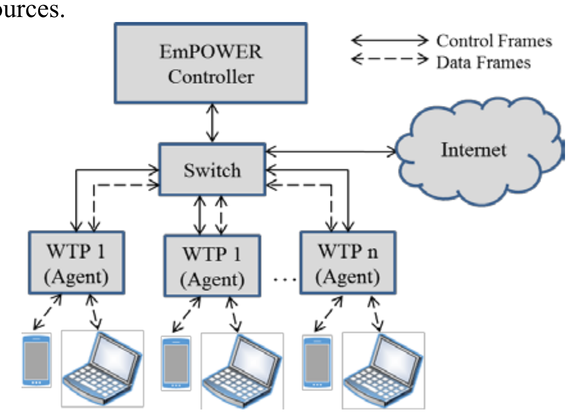

Figure 1: System Architecture

Each WTP hosts a radio agent as presented in Figure 1, which implements an interface for controlling WiFi-specific transmission settings and for gathering wireless related measurement data such as link and channel statistics. For the implementation of the agent, the Click Modular Router [4] has been used, which is a software implementation of a router that allows to be reconfigured by modifying already existed packet processing modules (known as elements) or by introducing new ones. The hypervisor is a Click element that belongs to the Agent configuration. It has to be pointed out that the hypervisor assigns dynamically the resources of each WTP in terms of

978-1-5090-6008-5/17/$31.00 ©2017 IEEE

allocated time, based on weights selected in the EmPOWER Controller.

III. Air time DRR (ADRR)

One of the challenges of a Wireless LAN lies on the fact that resources cannot be shared with respect to the assigned bandwidth, but must be shared in time. Thus, Round Robin (RR) or Deficit Round Robin (DRR) [5] could not be considered as candidates for the implementation of the hypervisor since they manage packets or transmitted bytes, respectively. As such, the proposed hypervisor uses a scheduling algorithm based on the principles of the Weighted DRR (WDRR), with major difference that instead of considering bytes for each queue, it considers the time necessary to transmit a packet. The logic behind the ADRR is as follows:

We start with a system quantum (Q_s) that is common for all the tenants and it corresponds to the time needed to transmit the packet of maximum size at the lowest bit rate. Then, for a given WTP, each tenant (i) is assigned a tenant quantum (Q_i) resulting by the multiplication of the Q_s with the agreed weight w_i (that reflects the desired sharing of the resources). Moreover, each tenant is assigned with a Deficit Counter (DC_i).

ADRR Algorithm

```
for ( i=0; i ≤ I; i ++;)
  initialize DC_i =0;
end for
for ( i=0; i ≤ I; i ++;)
  if (!queue.empty)
    DC_i = DC_i + Q_i;
    while (!queue.empty && DC_i ≥ 0)
      calculate_t_i,p();
      if (t_i,p ≤ DC_i)
          transmit packet;
          DC_i = DC_i – t_i,p
      end if
    end while
  end if
  else if (queue.empty)
    DCi = 0;
  end else if
end for
```

Figure 2: ADRR Algorithm

Starting from time t=0, the DC_i of each tenant is initialized to 0. Then, the algorithm checks if there is a packet to be transmitted. In case that there is a packet, it sets DC_i equal to the value of Q_i, it calculates the time needed for the transmission of it ($t_{i,p}$) and it checks if:

$$t_{i,p} \leq DC_i \qquad (1)$$

In the case that the above equation is true, the packet is transmitted and DC_i is set equal to DC_i - $t_{i,p}$. If it is false, the packet is not transmitted. In both cases, in the next iteration DC_i is incremented by Q_i. If the tested queue has no packets to transmit, then DC_i is set to zero and the next queue in line is checked. The pseudocode of the algorithm is presented in Figure 2. Let us notice that the calculation of $t_{i,p}$ is carried out based on an estimation of the bit rate that is going to be selected for the transmission of the next packet. This information can be drawn by the rate selection algorithm that is implemented as an element in the Agent (Click configuration).

IV. Demonstration

The proposed demonstration will show how the hypervisor can reconfigure the assigned resources to each tenant according to the selected weights. For that purpose we consider a scenario, depicted in Figure 3, consisting of one WTP and two tenants with one user each one. For the demonstration two different use cases will be considered based on the packet length for each tenant. In the first case we consider the same packet size (1500 bytes) for both tenants, while in the second case we consider packets of different sizes, i.e. 1500 bytes for tenant 1 and 500 bytes for tenant 2.

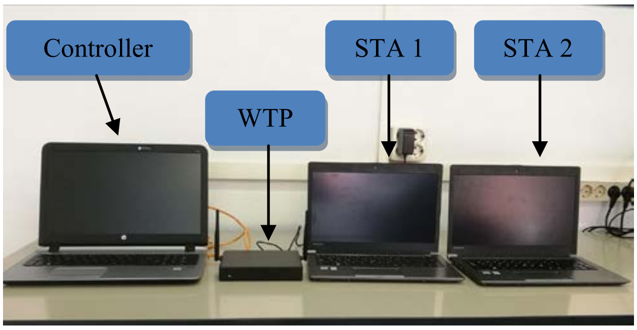

Figure 3: Demo platform

Using the above presented testbed the following will be demonstrated:

Resource allocation. In the first part of the demo we will target to demonstrate the benefits obtained from the use of the hypervisor employing the ADRR algorithm. It will be shown how the available resources are being shared in terms of assigned BW through the use of the online tool iperf.

Comparison with benchmark schemes. In the second part, a comparison with the WDRR will be carried out with purpose to depict the difference between the two scheduling algorithms and to explain the need for the proposed scheme.

Acknoledgements

This work has been supported by the RAMSES (Ref. TEC2013-41698-R), funded by CICYT (Spain). Moreover, we would like to express our keen gratitude to the Future Networks Research Unit (FuN) of the FBK CreateNet center (Italy) for providing us with the 5G-EmPOWER platform.

References

[1] H. Hawilo, *et al*, "NFV: state of the art, challenges, and implementation in next generation mobile networks (vEPC)," in *IEEE Network*, vol. 28, no. 6, pp. 18-26, Nov.-Dec. 2014

[2] Y. Li and M. Chen, "Software-Defined Network Function Virtualization: A Survey," in *IEEE Access*, vol. 3, no. , pp. 2542-2553,

[3] R. Riggio, *et al*, "Programming Abstractions for Software-Defined Wireless Networks," in *IEEE Transactions on Network and Service Management*, vol. 12, no. 2, pp. 146-162, June 2015.

[4] E. Kohler, *et al*, "The click modular router," ACM Transactions on Computer Systems (TOCS), vol. 18, no. 3, pp. 263–297, 2000

[5] M. Shreedhar and G. Varghese, "Efficient fair queuing using deficit round-robin," in *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 375-385, Jun 1996.