# Towards Wireless Infrastructure-as-a-Service (WIaaS) for 5G Software-Defined Cellular Systems

Albert Gran, Shih-Chun Lin, and Ian F. Akyildiz

Broadband Wireless Networking Lab, Georgia Institute of Technology, Atlanta, GA 30332, USA Email: alberto@ece.gatech.edu; slin88@ece.gatech.edu; ian@ece.gatech.edu

Abstract-As a key enabling technology for 5G cellular systems, wireless virtualization allows multiple service providers to simultaneously and independently serve their users via virtualized network slices. However, differently from wired network virtualization that has been studied for many years, the research of wireless resource slicing is still at a very early stage. In this paper, based on the proposed 5G software-defined systems, novel wireless infrastructure-as-a-service (WIaaS) is introduced, which enables mobile virtual network operators to provide distinguished services to their subscribed users while sharing a common physical infrastructure. Specifically, through softwaredefined networking and fine-grained base station designs, a throughput-efficient resource allocation is proposed, by which, at the same time, (1) the data-rate requirements of traffic flows in virtual networks are fulfilled, (2) the isolation among applications and deployed protocols in networks is guarded, and (3) the global resource utilization is maximized. Simulations confirm that the proposed solution outperforms state-of-the-art schemes with greater system throughput and fairness support. Moreover, the performance improvement becomes significant when transmitted data has real-time requirements or the flow density and diversity are increased. Thus, WIaaS facilitates wireless resource slicing upon software-defined architectures and has opened a new research area of virtualization in next-generation cellular systems.

## I. INTRODUCTION

Recently, the demands for higher data rates, lower end-toend latency, enhanced quality-of-service (QoS) for end-users, the exponential growth of multimedia applications, service diversity, and the radio access technology (RAT) heterogeneity have challenged current cellular system architectures to have a dramatic paradigm shift towards the forthcoming 5G next generation systems [1], [2]. It becomes more clear that a scalable, resilient, and flexible software-defined architecture with key technologies such as cloudification and virtualization will play a fundamental role. Among these enabling technologies, wireless virtualization has great potential to be exploited but is still at a very early stage of research. Specifically, by actively sharing the physical wireless infrastructure, it provides a powerful framework for high-level resource utilization, satisfying QoS requirements, and enhancing cross-technology functionalities in order to resolve the challenges of the next generation cellular systems. As a result, wireless virtualization significantly reduces capital and operational expenses, and facilitates new technology deployment by creating independent virtual networks, which simultaneously serve research testbeds and real operating networks [3], [4].



Fig. 1: Wireless Infrastructure as a Service (WIaaS) in 5G software-defined cellular systems.

Following the idea of wireless virtualization, in this paper, we propose innovative wireless infrastructure as a service (WIaaS) in 5G software-defined cellular systems. WIaaS allows different mobile virtual network operators to provide distinguished services to their subscribed users by sharing a common physical infrastructure. Via software-defined architecture, WIaaS introduces new dynamic wireless resource sharing, in which service providers lease the wireless infrastructure according to the instantaneous demands of their subscribed users. Therefore, the wireless hardware infrastructure is offered as a service rather than as a physical asset. There are two main advantages for WIaaS. First, service providers are given the ability to control, optimize, and customize the underlying infrastructure without owning it. They only pay for current used resources and do not need to predict future usages nor potential capacity peaks, leading to more costefficient operations. Second, isolation of resource allocations is controlled among service providers sharing the same infrastructure. Various networking capabilities required by different service providers can be integrated and deployed over the same



Fig. 2: SoftAir system architecture [5].

network infrastructure, while QoS requirements for all service providers are faced without interfering each other's operations and performance.

As indicated in Figure 1, we propose the realization of WIaaS through the design of a wireless hypervisor upon software-defined systems. Specifically, given the limited wireless resources, an efficient hypervisor adaptively virtualizes the physical wireless infrastructure by determining optimal resource allocation at each moment. With the allocated virtualized resources, service provider users can receive or transmit the desired data until the next decision is performed. Therefore, the hypervisor (1) supports efficient allocation for the limited physical resources, (2) facilitates service providers to accommodate the dynamic demands of their subscribed users in virtual networks, and (3) enables resource slicing for users' flows under intra-slice customization and inter-slice isolation. Thus, the designated wireless hypervisor serves as a practical and efficient solution to carry out wireless virtualization in 5G cellular systems.

To the best of our knowledge, this work is the first to propose throughput-efficient, utilization-optimal wireless virtualization at a flow-level granularity. The rest of the paper is organized as follows. Section II introduces software-defined cellular architecture that brings the separation between the control plane and the programmable data plane. To enable dynamic resource allocation with respect to a great variety of service requirements in traffic flows, Section III provides the proposed wireless hypervisor that simultaneously maximizes resource-utilization, ensures serving fairness among application flows, and controls transmit power to enhance signalto-noise-ratio (SNR) requirements. Section IV presents the performance evaluation and Section V concludes the paper.

#### II. 5G SOFTWARE-DEFINED CELLULAR SYSTEMS

#### A. System Model

The flexibility required for a granular and efficient network virtualization can only be achieved with a radical paradigm shift in current cellular architectures. For instance, a software-defined based architecture, called SoftAir, is proposed in [5], offering all the essential functionalities for 5G wireless communication systems. The software-defined architecture decou-

ples the data plane from the control plane and facilitates the implementation of network virtualization functionalities. In the data plane, a software-defined radio access network (SD-RAN) composes of fine-grained distributed software-defined base stations (SD-BSs), which allows the separation of the baseband processing and the MAC layer operations at the baseband server (BBS) from the radio frontend implementation at the remote radio heads (RRHs). On the other hand, in the control plane, through the software coordination from a cloud environment, an efficient and adaptive sharing of network resources can be designed to maximize spectrum efficiency and allow the convergence of heterogeneous networks.

As shown in Figure 2, the network virtualization for SD-RANs is achieved through the wireless hypervisor. It is a low-level resource scheduler in the SD-BSs that executes the optimal resource allocation for each virtual network, regarding channel conditions and RAT via a variety of wireless resource dimensioning schemes. Moreover, all the resource-allocation policies executed in the wireless hypervisor are decided by the network hypervisor at the network controller, which takes advantage of the whole view of the network. As a result, both wireless and network hypervisors, together with a switch hypervisor in the software-defined switches in the core network, provide SoftAir with the capability of enabling end-to-end network virtualization.

## B. Wireless Virtualization: Requirements and Key Designs

The objective is to determine the allocation of nonconflicting network resource blocks via wireless hypervisor among virtual network operators based on their demands and maximizing the global resource utilization. However, coexistence of different service providers sharing the physical infrastructure is not the only requirement for an accurate virtualization. Each service provider has to be given freedom to customize their virtual networks in terms of different topologies, number of users, data rates and QoS requirements, and moreover, any change of these parameters in one of the virtual networks can not affect the behavior or performance of the rest of the virtual networks, i.e. the different virtual networks should be transparent from each other.

When considering wireless virtualization, channel-aware mechanisms should be provided with the flexibility and isolation requirements for resource sharing. Specifically, the only way to achieve high spectral efficiencies when performing the scheduling is to give higher priorities to flows with better channel conditions, taking advantage of the radio resource management (RRM) techniques that current cellular systems contain (i.e. CQI reporting). Although the main objective is always the satisfaction of data rate QoS requirements of different virtual network flows, fairness should also be controlled at a slice, flow and user granularity, to allow different protocols and applications be ran together without affecting other's performance. Providing backwards compatibility, which implies developing a solution that do not require an extensive modification of the MAC/PHY layers, is another interesting challenge if an easy-to-be-deployed solution is being looked



Fig. 3: Wireless resource virtualization scenario.

after. The scheduling has to be performed according to the existing specifications. In current cellular systems, schedulers work with a granularity of a certain transmission time interval (TTI) and a specific spectrum bandwidth, e.g. 1ms and 180kHz for a resource block (RB) in LTE. Finally, although computational complexity is also one of the strongest requirements, software-defined networking architectures allow the usage of high power computation by running scheduling algorithms in network controllers, located in data centers.

## III. RESOURCE-EFFICIENT WIRELESS VIRTUALIZATION WITH FAIRNESS SUPPORT

## A. Problem Description

Consider a SD-BS in a SD-RAN, which consists of a single BBS and a RRH. Assume that users are randomly distributed and served by the SD-BS with different channel conditions. These users subscribe services from the virtual networks, created by the SD-BS. Specifically, a single virtual network can offer a highly diverse range of services to users by setting up the corresponding QoS requirements to the SD-BS. Moreover, every association in a virtual network where a user receives services from is defined as flow. That is, a flow characterizes a group of data from a specific virtual network service with dedicated user and QoS requirements.

To forward different flows to users, the BS needs to allocate the physical resources among flows. Assuming that resource allocation is performed in terms of RBs and the decisions are taken every specific TTI, at each TTI the BS has to assign one or several resource blocks to each flow [6]. Note that only downlink is considered. The objective is to find the optimal allocation of available RBs to different flows of the BS in an efficient way (i.e. taking the maximum profit of the available resources) and strictly satisfying the data rate QoS requirements specified by the flows. The fundamental point is that the allocation of these RBs does not affect the same way if they are allocated to one flow or another. Depending on user instantaneous channel conditions of each flow, that flow will be able to transmit more or less information in the same RB.

Two parameters are considered for the profit: (1) the throughput required that each flow needs to transmit to satisfy its data-rate requirements and (2) the throughput provided by the channel that depends on the channel conditions of the user. Comparing these two parameters, the best allocation is the one that selects the flows taking the most profit of the channel at each moment, maximizing resource-utilization and satisfying at the same time the data rate requirements of the flows by checking if the throughput provided is bigger than the required. The second concept defined comes from the fact that only taking into account the profit one can find situations in which all the resources are allocated to specific flows, users, service providers or virtual networks with high peaks of profit, not allowing the other instances of the network to transmit. Therefore, we consider the fairness, which controls the system isolation at three levels: slice, flow and user, while providing the enough flexibility to mobile network operators for determining the differentiation policy among operators. Finally, a power re-allocation is considered to enhance global signal-to-noise ratio and minimize energy consumption.

## B. Throughput-Efficient Resource Allocation for WIaaS

We consider WIaaS with respect to the profit, the fairness, and the available power at a SD-BS in the following.

1) *Profit:* Assume each flow in a BS has two QoS requirements. The first is the required throughput, defined as:

$$R_f = Arrival\_rate_f \cdot TTI; \tag{1}$$

$$Th\_required_f = \frac{R_f}{QoSDelay_{max,f}},$$
(2)

where  $R_f$  is the amount of data that flow f has to transmit in the next scheduling decision according to its specific  $Arrival\_rate$  and inter-decision interval time (TTI), and the  $QoSDelay_{max,f}$  refers to the maximum time that flow f data can stand before reaching the user. The satisfaction of this requirement depends basically on the number of RBs that are allocated to the flow, and the amount of information that can be carried in each of those RBs.

To consider full RRM techniques availability, it is assumed that channel quality information (CQI) reports from all users are received periodically by the BS. The CQI reports show the information that can be transmitted according to the channel state at every moment. The amount of data that flow f from user u can transmit from the entire RB bandwidth in one scheduling decision is:

$$Th\_provided_f \cdot TTI = \sum_{r=1}^{RB} i_{r,f} \cdot Data\_provided_{r,f}$$
  
s.t. 
$$\sum_{f \in F} \sum_{r \in RB} i_{r,f} \leq RB\_BW,$$
 (3)

where  $i_{r,f} = 1$  when RB *r* is assigned to flow *f* in the current decision and 0 otherwise.  $Data\_provided_{r,f}$  measures the amount of flow *f* information that RB *r* can carry according to the code rate and modulation scheme selected from the last CQI feedback information. In this way, the profit that the flow is taking from the provided resources can be defined as:

$$PF_{f}(\%) = \begin{cases} \frac{Th\_required_{f}}{Th\_provided_{f}} \cdot 100\\ 0 \text{ if } \sum_{r=1}^{RB} i_{r,f} = 0 \end{cases}$$
(4)

As the most efficient solution is the one allocating RBs to the flows taking the most profit from the channel limited resources at each scheduling decision, given F the set of flows allocated in the BS, we formulate the problem as follows:

$$\begin{array}{ll}
\max_{i_{r,f}} & \frac{1}{F} \sum_{f \in F} PF_f \\
\text{s.t.} & PF_f \le 100, \quad \forall f \in F.
\end{array}$$
(5)

2) Fairness: One of the main requirements for wireless virtualization is isolation, which implies that all the customization changes realized in one of the virtual networks in terms of topology, number of users or even data rate requirements can not interfere the behavior of its neighbors. Although achieving complete isolation is not possible, at least some fairness control should be developed, allowing all the different virtual networks to transmit even under bad channel conditions.

The flow fairness is defined as follows. A historical record of the profit obtained by each flow, called Fairness Index (FI) is stored and updated at every scheduling decision, allowing the BS to know how much utilization of the channel resources has the flow been achieving. In this way, if in the next decision a specific flow is selected to transmit, the new profit obtained by the flow will be averaged with the historical value stored, dynamically updating the fairness parameter. If in a specific moment this fairness index is decreasing too much, more priority can be given to that flow when performing the next scheduling decision. In the same way, if the fairness index of a flow is increasing too much, it may be convenient to give priority to the other flows, even though their instantaneous profits are smaller, providing equity and opportunities between all the flows and controlling the system isolation. Therefore, we have

$$FI_{f}(d) = \frac{FI_{f}(d-1) \cdot D_{d-1} + PF_{f}(d)}{D_{d}}, \quad (6)$$

where  $D_d$  and  $D_{d-1}$  stand for the number of scheduling decisions taken until the *d* and *d-1* instant respectively, since the flow is assigned to the BS. For the fairness control, a new constraint in the optimization should be taken into account:

$$\varphi_1 \le FI_f \le \varphi_2, \quad \forall f \in F.$$
 (7)

When sharing resources among flows in the network, the flow fairness should be considered with the following reason. It is very common to have situations in which different parties impact on the network needs to be controlled, not only at a flow level, but also at a virtual network level or even at a user level. In literature a lot of proposals consider scenarios in which the different virtual networks are treated differently based on the distinguished possible contracts with MNOs [?], [7] or based on the differentiation between the infrastructure owner and its tenants [8]. One may also need to provide equity between users of a virtual network in terms of transmission opportunities when allocating the resources. All these scenarios depend on the operator policies, but a completely flexible solution exploiting the fairness concept can be implemented by broadening the idea of fairness at three levels: flow (as we had), but also virtual network (or slice) and user. The fairness index has to be computed following the same methodology for the virtual networks and the users, with their corresponding flows information. Given  $F_u$  and  $F_s$ as the flow set of user u and slice s, respectively, we have

$$i_f = 0$$
 when  $\sum_{r=1}^{RB} i_{r,f} = 0$ ,  $i_f = 1$  otherwise; (8)

$$PF_u = \frac{1}{F_u} \sum_{f \in F_u} i_f \cdot PF_f; \tag{9}$$

$$PF_s = \frac{1}{F_s} \sum_{f \in F_s} i_f \cdot PF_f.$$
<sup>(10)</sup>

Specifically, for all users, we have

$$FI_{u}(d) = \frac{FI_{u}(d-1) \cdot D_{d-1} + PF_{u}(d)}{D_{d}}; \qquad (11)$$

and for all slices, we have

Ì

$$FI_{s}(d) = \frac{FI_{s}(d-1) \cdot D_{d-1} + PF_{s}(d)}{D_{d}}.$$
 (12)

3) Power: Exploiting channel conditions, data rate requirements and power available of the base-station, a re-allocation of power can be implemented for the satisfaction of not only the data rate, but also the signal-to-noise ratio requirements of the different flows. The basic idea is to use the power assigned to the not-being-used RBs, taking into account that LTE distributes the available power uniformly among all the RB-bandwidth, to achieve higher levels of quality in the RBs used for transmission. The iteration goes as follows: (1) compute the maximum throughput that can be achieved with a uniformly distributed power; (2) calculate the needed power to achieve a more strict quality for the same throughput; (3) redistribute the power of the non-used RBs among the used ones maximizing the quality requirements satisfied.

To achieve a certain  $(Eb/No)_{min,f}$  in reception, for a specific throughput already established, the BS needs to transmit the signal at a specific power. This power depends on the prerequired quality, the channel conditions, the antenna gains, and the path loss. That is,

$$P_{T\_required_f} = \frac{Th\_required_f\left(\frac{Eb}{No}\right)_{min,f}\left(\frac{P_N}{B}\right)L_u}{G_T G_R}$$
(13)

where  $G_T$ ,  $G_R$ ,  $P_N$ , B and  $L_u$  are the transmission and reception antenna gains, the noise power, the bandwidth and the propagation losses of the channel including both interference and fading margins, respectively.

From the assumption of full RRM mechanisms availability, including link adaptation, the channel instantaneous conditions can be obtained as follows. CQI in LTE gives the maximum modulation and code rate that can be used according to the channel conditions, so that a block error rate of 0.1 is maintained at the UE [9]. Therefore, the operation performed for the CQI computation, can be mathematically described as:

$$Rb_{\max} = rmB = \frac{P_{T\_CQI}G_TG_R}{\left(\frac{P_N}{B}\right)\left(\frac{Eb}{No}\right)_{\min\_CQI}L_u}, \qquad (14)$$

where code rate r and modulation m are found according to the LTE uniformly distributed power  $P_{T\_CQI} = P_{eNB}/(RB)$ , B = 12 sub-carriers of 15000 Hz for a single RB, and  $\left(\frac{Eb}{No}\right)_{\min\_CQI}$  the minimum to achieve a BLER < 10% for the current code and modulation scheme. In this way, we have

$$\frac{G_T \cdot G_R}{\left(\frac{P_N}{B}\right) \cdot L_u} = \frac{r \cdot m \cdot B \cdot \left(\frac{Eo}{No}\right)_{\min\_CQI}}{P_T\_CQI};$$
(15)

Thus, to achieve a more strict Eb/No requirement for the same throughput, code rate, and modulation, we have

$$P_{T\_required_f} = \frac{Th\_required_f \left(\frac{Eb}{No}\right)_{min,f} P_{T\_CQI}}{rmB\left(\frac{Eb}{No}\right)_{min\_CQI}};$$
(16)

$$P_T\_total = \sum_{f \in F} i_f \cdot P_T\_required_f \tag{17}$$

where all the parameters are known. Power allocation can suppose an important improvement if SNR requirements are desired to be satisfied. Verifying if the power available in the BS for a specific radio access technology is enough for transmitting the amount of data required with the energy needed to satisfy their demanded quality, will be reflected as another constraint bounding the obtained results. Thus, we formulate the optimization problem as follows:

$$\max_{\substack{i_{r,f} \\ s.t.}} \sum_{\substack{f \in F \\ f \leq 100 \\ \varphi_1 \leq FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{\substack{f,s \\ FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in F \\ f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in F \\ f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in F \\ f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in F \\ f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \leq Total\_power_{BS} }} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \geq P}} \sum_{j=1}^{N} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \geq P}} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \leq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI_{u,f,s} \geq \varphi_2 \\ P_T\_total \geq P} \sum_{\substack{f \in FI$$

The formulated optimization problem can be solved by commercial solvers through exhaustive searching methods. The acquired solutions are then executed by wireless hypervisors for the optimal resource sharing policies. Therefore, the proposed WIaaS adaptively computes the resource allocation with respect to available bandwidths at each time and facilitates the dynamic deployment of virtual networks and their flows.

#### **IV. PERFORMANCE EVALUATION**

In this section, based on a developed Java simulator, we present simulation results to evaluate the performance achieved by WIaaS in a realistic wireless setup. In particular, the system throughput and fairness are examined in detail as follows.

#### A. Throughput-efficiency Assessment

We first evaluate the achievable throughput by our proposed solution (MP) and compare it to existing designs that use maximum signal-to-noise ratio (MSNR) and weighted fair queuing (WFQ) algorithms. As the conventional MSNR scheme is user-based, allocating the available resources to the user with best channel conditions, a flexible adaptation of the algorithm has been created to provide flow-level granularity, in which flow combinations are the ones compared, and the best one is selected at each scheduling decision. For the WFQ in wireless communications, the channel rate is different for all the users and the resource partition is conditioned by the RB size. For

Flow	User	VN	Service	Arrival Rate	QoS Delay
1	1	1	HDTV	30000 kbps	40 ms
2	1	2	File Transfer	6000 kbps	40 ms
3	1	3	VoIP	80 kbps	20 ms
4	2	1	HDTV	30000 kbps	40 ms
5	2	1	File Transfer	6000 kbps	25 ms
6	2	2	Web Browsing	5000 kbps	30 ms
7	3	1	RT Gaming	10000 kbps	15 ms
8	3	3	VoIP	80 kbps	20 ms
9	3	3	VideoCall	1000 kbps	25 ms

TABLE I: Simulation Parameter Setting.

this reason, an adaptation of the WFQ has also been considered in which the weight of each flow is the fraction relating flow throughput required with the sum of all flow throughput demands. To match the RB size, a progressive recursion is defined decreasing the total weight until all the available RBs can be filled by different ows according to their weight priority. TABLE I summarizes the simulation parameter setting, which has been selected following the LTE quality class information defined by 3GPP [9]. From the different simulations, two parameters are analyzed: the achieved throughput for total data amount transmitted as a result of the scheduling decision, and the useful throughput achieved for the data amount with regard of QoS requirements.

The results for scheduling algorithms of WFQ and MSNR are shown in Figure 4. One can clearly see that maximum SNR technique is not efficient, as having good channel conditions does not necessarily implies having a lot of data to transmit. In the same direction, maximum SNR is almost the opposite of a fair algorithm, discarding the users with bad channel conditions regardless of their arrival rates, queues or requirements of their services and not providing isolation, which is the most crucial point for virtualization. In Figure 4, despite that very high throughput has been obtained and all flows have enough data to transmit, it shows that the big majority of data is not satisfying its QoS delay requirements. For the WFQ the results are very similar, although good throughput is achieved, absolutely none of it is useful, due to the fact that, giving just a fraction of what every flow needs, although making the algorithm fair, is not enough for the requirement satisfaction. This implies that the transmitted data will be discarded in reception, making the algorithm inefficient in terms of throughput and power. On the other hand, our algorithm achieves higher throughput than other schemes, facilitating efficient resource allocation.

## B. Fairness Scheme Analysis

The second objective of our hypervisor solution is to provide isolation between different virtual networks. Figure 5 shows the evolution of the fairness index at a virtual network level. In the upper sub-figure, no fairness control has been implemented. It shows that, while VN1 is achieving high profit levels, VN2, and more specially VN3, remain with very low values of historical profit achieved. One option for this situation could be applying a fairness control just to VN3, allowing it to transmit although having low profit levels. However, these



Fig. 4: Throughput for different scheduling algorithms.

adjustments would depend on the specific situation, which is not an ideal solution, since an analysis has to be performed in advance to determine a correct policy for every scenario. To manage the fairness in an adaptive and dynamic way, working with the standard deviation parameter is proposed. It implies quantifying the amount of variation or dispersion of the different entities (e.g., virtual networks, users or flows) fairness indexes and taking scheduling decisions to restrict its value to be lower than a certain bound. Although the operator is given maximum flexibility to develop customized policies, it can be seen that the presented technique is an appropriate solution for achieving efficient and dynamic fairness-control.

## V. CONCLUSION

This paper introduces WIaaS in software-defined cellular systems by developing a throughput-efficient wireless hypervisor. The objective is to determine the optimal allocation of non-conflicting wireless resource blocks among virtual network operators, so that the data-rate requirements are satisfied and the global throughput efficiency is maximized at the same time. Through a mathematical optimization framework, the proposed hypervisor captures different traffic requirements and channel conditions at every moment and enables infrastructure sharing at levels of slices, flows, and users. Based on a highly manageable and configurable simulator, performance evaluation confirms the optimality of our solution with considerable performance improvement than existing schemes, in a variety of wireless environment setups. With the presented work, we have opened a new research area for centralized control to

Fairness Index Evolution 5 Simulations - 2RB No Fairness Control



Fig. 5: Fairness among virtual networks.

tackle virtualization challenges in next generation 5G cellular systems.

## ACKNOWLEDGEMENT

This work was supported by the US National Science Foundation (NSF) under Grant No. 1547353.

#### REFERENCES

- J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *Selected Areas in Communications*, *IEEE Journal on*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [2] I. F. Akyildiz, S. Nie, S.-C. Lin, and M. Chandrasekaran, "5G roadmap: 10 key enabling technologies," *Computer Networks*, vol. 106, pp. 17–48, 2016.
- [3] C. Liang and F. R. Yu, "Wireless network virtualization: A survey, some research issues and challenges," *Communications Surveys & Tutorials*, *IEEE*, vol. 17, no. 1, pp. 358–380, 2015.
- [4] I. F. Akyildiz, S.-C. Lin, and P. Wang, "Wireless software-defined networks (W-SDNs) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation," *Computer Networks*, vol. 93, Part 1, pp. 66–79, 2015.
- [5] I. F. Akyildiz, P. Wang, and S. C. Lin, "Softair: A software defined networking architecture for 5G wireless systems," *Computer Networks*, vol. 85, pp. 1–18, Jul. 2015.
- [6] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.
- [7] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, "A novel LTE wireless virtualization framework," in *Mobile Networks and Management*. Springer, 2010, pp. 245–257.
- [8] X. Lu, K. Yang, Y. Liu, D. Zhou, and S. Liu, "An elastic resource allocation algorithm enabling wireless network virtualization," *Wireless Communications and Mobile Computing*, vol. 15, no. 2, pp. 295–308, 2015.
- [9] S. Sesia, I. Toufik, and M. Baker, *LTE-the UMTS long term evolution*. Wiley Online Library, 2015.