SDN-Based Smart Gateways (Sm-GWs) for Multi-Operator Small Cell Network Management

Akhilesh S. Thyagaturu, Yousef Dashti, and Martin Reisslein, *Fellow, IEEE*

***Abstract*—Small wireless cells have the potential to overcome bottlenecks in wireless access through the sharing of spectrum resources. However, current backhaul networks statically allocate resources, resulting in minuscule allocations when many small cells are connected to cellular operators with given resources. We introduce a novel access backhaul network architecture based on a smart gateway (Sm-GW) between the small cell base stations, e.g., long term evolution (LTE) enhanced Nodes B (eNBs), and the conventional backhaul gateways, e.g., LTE servicing/packet gateways (S/P-GWs). We specify the modest LTE protocol modifications that integrate the Sm-GW into the conventional LTE network. The Sm-GW flexibly schedules uplink transmissions for the eNBs. Our simulation evaluations indicate that the Sm-GW scheduling can fairly allocate uplink transmission bitrates to the eNBs and reduce packet delays. Based on software defined networking (SDN), we introduce a management mechanism that allows multiple operators, i.e., multiple S/P-GWs, to flexibly inter-operate via multiple Sm-GWs with a multitude of small cells. An SDN orchestrator coordinates the adaptive allocation of uplink transmission bitrates to Sm-GWs (which in turn allocate the uplink transmission bitrates to eNBs based on their demands). We formulate optimization problems for the operator (S/P-GW) resource allocation to Sm-GWs without and with sharing among operators. Our numerical evaluations indicate that the flexible SDN orchestration substantially increases the network throughput compared to the current static resource allocations.**

***Index Terms*—Backhaul, multi-operator network management, small cells, software defined networking (SDN).**

I. Introduction

A. Motivation: Small Cells

RECENT wireless communications research has examined the benefits of splitting the conventional cells in wireless cellular communications into small cells for supporting the growing wireless network traffic. Small cells can coexist with neighboring small cells while sharing the same spectrum resources, and are thus an important potential strategy for accommodating wireless network traffic growth [1]. Small cells are also sometimes referred to as “femto” cells in the context of the Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) wireless standard; we use the general terminology “small” cells throughout. However, small cells pose new challenges, including interference coordination [2], backhaul complexity [3], [4], and increased network infrastructure cost [5]. In this article we propose a solution to reduce the infrastructure cost and complexity of backhaul access networks supporting small cells.

Small cell networks are expected to be privatively owned [6]. Therefore it is important to enable usage flexibility and the freedom of investment in the new network entities (e.g., gateways and servers) and the network infrastructures (e.g., switches and optical fiber) by the private owners of small cells. While a plethora of studies has examined advanced enhanced Node B (eNB) resource management, e.g., [7]–[9], the implications of small cell deployments for backhaul gateways have largely remained unexplored [10]. Generally, backhaul access networks that interconnect small cell deployments with LTE gateways can employ a wide variety of link layer (L2) technologies, including SONET/SDH, native Ethernet, and Ethernet over MPLS [11]–[13]. In order to accommodate these heterogeneous L2 technologies, cellular LTE network interfaces, such as S1 and X2 interfaces, are purposefully made independent of the L2 technology between small cell deployments and gateways. Due to the independent nature of L2 technologies, a dedicated link with prescribed QoS, which can support the fundamental operations of cellular protocols, must be established for each interface connection [14]. Statistical multiplexing is then limited by the aggregate of the prescribed QoS requirements and only long-term re-configurations, e.g., in response to deployment changes, can optimize the backhaul transmissions [15]. Present wireless network deployments based on the 3GPP LTE standard do not provide feedback from the eNBs to a central decision entity, e.g., an SDN orchestrator, which could flexibly allocate network resources based on eNB traffic demands. Thus, present wireless backhaul architectures are characterized by (*i*) essentially static network resource allocations between eNBs and operator gateways, e.g., LTE Servicing/Packet Gateways (S/P-GWs), and (*ii*) lack of coordination between the eNBs and the operator gateways in allocating these network resources, resulting in under-utilization of the backhaul transmission resources. Additionally, exhaustion of available ports at

Manuscript received March 15, 2016; revised June 28, 2016 and August 26, 2016; accepted August 31, 2016. Date of publication September 2, 2016; date of current version December 8, 2016. The associate editor coordinating the review of this paper and approving it for publication was V. Fodor. *(Corresponding author: Martin Reisslein.)*

A. Thyagaturu and M. Reisslein are with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: athyagat@asu.edu; reisslein@asu.edu).

Y. Dashti is with the Electronics Engineering Technology Department, College of Technological Studies, Public Authority for Applied Education and Training (PAAET), P.O. Box 42325, Shuwaikh 70654, Kuwait (e-mail: yy.dashti@paaet.edu.kw).

Digital Object Identifier 10.1109/TNSM.2016.2605924

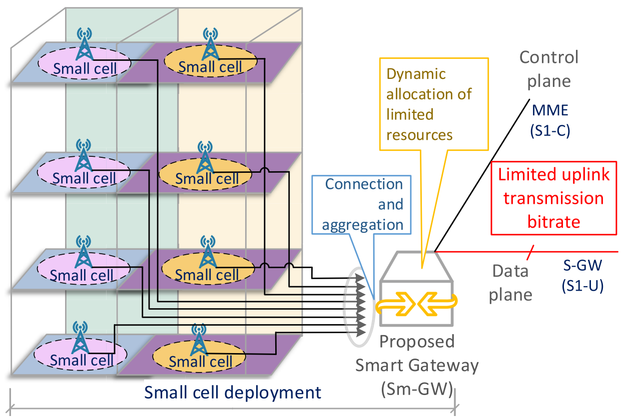


Fig. 1. The proposed smart gateway (Sm-GW) is inserted between the small cell base stations (LTE eNBs and access points) and the conventional operator core network entities, such as LTE Mobility Management Entity (MME) and Servicing Gateway (S-GW). The Sm-GW flexibly aggregates the connections to a multitude of small cell base stations and dynamically allocates the limited uplink transmission bitrate to the different small cells.

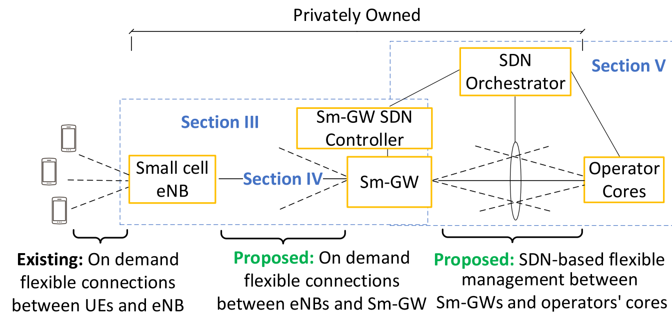


Fig. 2. The Smart Gateway (Sm-GW) architecture, see Section III, consists of privately owned Sm-GWs that are inserted between the existing conventional operator core (e.g., LTE S/P-GWs) and small cell eNB base stations. Flexible Sm-GW scheduling (Section IV) assigns uplink transmission bitrates on demand to the small cell base stations. The uplink transmission bitrates from multiple operators are optimally assigned to the Sm-GWs by an SDN orchestrator (Section V).

the operator gateways can limit the eNB deployment in practice.

The static resource allocations and lack of eNB-gateway cooperation are highly problematic since the aggregate uplink transmission bitrate of the small cells within a small geographic area, e.g., in a building, is typically much higher than the uplink transmission bitrate available from the cellular operators. Thus, small cell deployments create a bottleneck between the eNBs and the operator gateways. For instance, consider the deployment of 100 small cells in a building, whereby each small cell supports 1 Gbps uplink transmission bitrate. Either each small cell can be allocated only one hundredth of the operator bitrate for this building or the operator would need to install 100 Gbps uplink transmission bitrate for this single building, which would require cost-prohibitive operator gateway installations for an organization with several buildings in a small geographical area. However, the uplink transmissions from the widespread data communication applications consist typically of short high-bitrate bursts, e.g., 100 Mbps bursts. If typically no more than ten small cells burst simultaneously, then the eNBs can dynamically share a 1 Gbps operator uplink transmission bitrate. An additional problem is that with the typically limited port counts on operator gateways, connections to many new small cells may require new operator gateway installations. An intermediate Sm-GW can aggregate the small cell connections and thus keep the required port count at operator gateways low.

B. Overview of Network Management With SDN-Based Sm-GW

We present a new backhaul network framework for supporting small cell deployments based on a new network entity, the Smart GateWay (Sm-GW). Consider an exemplary small cell deployment throughout multiple buildings of a university. Each building has hundreds of small cells that are flexibly connected to an Sm-GW, as illustrated in Fig. 1. Multiple Sm-GWs are then connected to core networks, i.e., the S-GWs and P-GWs, of multiple cellular operators via physical links (e.g., optical or microwave links), as illustrated for a single Sm-GW in Fig. 2. An SDN orchestrator owned by the university manages the cellular infrastructure of the entire university. The SDN orchestrator coordinates the resource allocations from the operators to the Sm-GWs.

The main original contributions of this article are:

1) A novel comprehensive Smart Gateway (Sm-GW) architecture and protocol framework that accommodates a flexible number of eNBs while reducing the requirements at the operator's core, e.g., at LTE S-GW and MME. The Sm-GW physically and logically aggregates the eNB connections so that a set of eNBs appears as a single virtual eNB to the operator gateways, see Section III.
2) A Sm-GW scheduling framework to flexibly share the limited uplink transmission bitrate among all the small cell eNBs connected to an Sm-GW, see Section IV.
3) An adaptive SDN-based multi-operator management framework that dynamically shares the uplink transmission bitrates of multiple operators among the Sm-GWs. An SDN orchestrator dynamically coordinates the sharing among the Sm-GWs, the transport network connecting the Sm-GWs to the operator gateways, and the operator gateways, see Section V.

C. Related Work

Recently proposed SDN based backhaul architectures, such as CROWD [16], [17], iJOIN [18], U-WN [19], Xhaul [20], the multi-tiered SDN based backhaul architecture [21], and similar architectures [22]–[27], are revolutionary designs proposing new cellular infrastructure installations. In contrast, our proposed SDN-based Sm-GW enables the softwarization of *existing* cellular infrastructures consisting of eNBs and conventional operator gateways, such as the S/P-GW in LTE core networks. The proposed Sm-GW is inserted in the existing backhaul infrastructure to inter-network and coexist with the existing LTE network core entities, such as the S/P-GW.

SDN based backhaul architectures with centralized interference coordination have been proposed in [28]–[30].

The SDN controller in these architectures maintains a global database of spectrum resources and dynamically assigns the resources to base stations so as to minimize the mutual interference among base stations in dense deployments. We note that wireless interference is a localized phenomenon. Therefore, a base station is most affected by its neighboring base stations in dense deployments. The centralized interference coordination techniques in [28]–[30] are complementary to our proposed Sm-GW architecture in that they can be implemented at the Sm-GW instead of the SDN controller/orchestrator.

Schedulers at the eNB allow multiple user equipment (UE) devices to share the wireless resources at the eNB. For example, the LTE standard medium access control (MAC) protocol [31] coordinates the scheduling of wireless resources between an eNB and multiple UEs. Generally, most wireless resource scheduling studies to date have focused on the sharing of the wireless resources at a given single eNB. For instance, quality of service (QoS) aware uplink scheduling and resource allocation at a given single small cell eNB in an LTE network have been examined in [32]. In contrast, we propose a novel scheduling framework at the Sm-GW based on uplink transmission bitrate requests from *multiple eNBs*, i.e., we propose the sharing of the backhaul network resources among multiple eNBs.

A similar sharing of network resources among small cell base stations has been studied in [33]. Specifically, the H-infinity scheduler for limited capacity backhaul links [33] schedules the traffic in the downlink. The centralized H-infinity scheduler focused on buffer size requirements at the base stations in the small cell networks. In contrast, we focus on the *uplink* traffic from the eNBs to the Sm-GW. To the best of our knowledge, we propose the first network protocol framework for the uplink transmissions from multiple eNBs to the operator gateways in the context of LTE small cells. We note that our Sm-GW framework is complementary to several recently studied resource allocation mechanisms in cellular networks. For instance, D2D resource allocation through traffic offloading to small cell networks has been studied in [34]; this D2D approach can be readily supported by our proposed Sm-GW. Coordinated scheduling in the context of small cells with dynamic cell muting to mitigate the interference has been discussed in [35]. The cell muting technique can be further extended based on our approach of traffic scheduling to eNBs. A flexible wireless resource allocation mechanism based on the SDN programmability of traffic flows from a single UE device to multiple base stations in dense small cell networks has been examined in [36]. The offloading of UE traffic for efficient traffic management in small cell networks has been examined in [37]. In contrast, we propose an SDN-based multi-operator resource allocation mechanism that allocates limited backhaul link capacities to multiple Sm-GWs (which in turn can flexibly allocate the capacities to multiple eNBs). The UE to eNB communication approach from [36] and UE traffic offloading [37] are thus complementary to our eNB to Sm-GW and Sm-GW to S/P GW network management approaches.

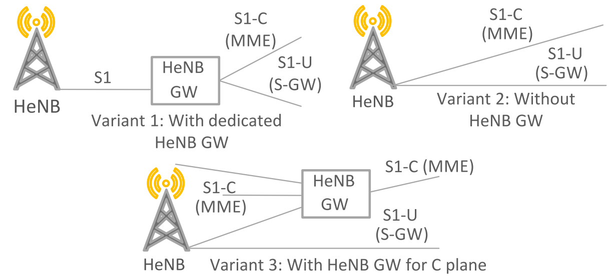


Fig. 3. HeNB architectural models in 3GPP LTE.

II. Background: Conventional LTE Small Cell Backhaul

In this section we describe the conventional architectural model for Home-eNodeB (HeNB) access networks [38] and the network sharing mechanism in 3GPP LTE. HeNBs are the small cell base stations of the LTE standard. We use the general terminology "eNB" to denote all types of small cell base stations.

A. HeNB Architectural Models in 3GPP LTE

In Figure 3 we show the 3GPP HeNB architectural models: 1) with dedicated HeNB-GateWay (HeNB-GW), 2) without HeNB-GW, and 3) with HeNB-GW for the control plane.

1) With Dedicated HeNB-GW: With a dedicated HeNB-GW, communication between the HeNB and the HeNB-GW is secured by a mandatory security gateway (Se-GW) network function. The HeNB-GW aggregates the control plane connections (S1-MME) and user plane connections (S1-U) of all HeNBs connected to the HeNB-GW to a single control and user plane connection. The HeNB gateway appears as a single eNB to the outside entities, such as S-GW and MME. In a similar way, the HeNB-GW appears as both an S-GW and an MME to the eNBs connected to the HeNB-GW. The numbers of ports required at the MME and S-GW are reduced through the aggregation at the HeNB-GW. Our proposed Sm-GW architecture is similar to the dedicated HeNB-GW architecture, in that the Sm-GW aggregates the eNBs connections both physically and logically. In addition, our Sm-GW flexibly allocates uplink transmission bitrates to small cell eNBs (see Section IV) and allows for the adaptive allocation of operator uplink transmission bitrates to the Sm-GW by the SDN orchestrator (see Section V).

2) Without HeNB-GW: Deployments of HeNBs without the HeNB-GWs increase the requirements on the S-GW and MME to support large numbers of connections. Large deployments of small cells without gateway aggregation at the HeNBs would greatly increase the total network infrastructure cost.

3) With HeNB-GW for the Control Plane: HeNB control plane connections are terminated at the HeNB-GW and a single control plane connection is established from the HeNB gateway to the MME. Although the number of connections required at the MME is reduced due to the control plane aggregation at the HeNB-GW, data plane connections are still terminated directly at the S-GW, increasing requirements at the S-GW. The Se-GW typically secures the communication

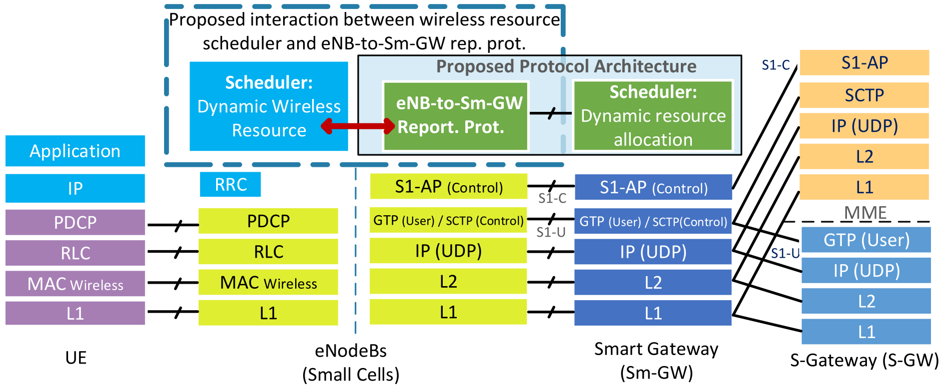


Fig. 4. Illustration of proposed protocol mechanisms at eNB, Sm-GW, and S-GW. At the eNB, the eNB-to-Sm-GW reporting protocol formulates the uplink transmission bitrate requests based on the UE traffic requirements (obtained through interactions with the wireless resource scheduler). At the Sm-GW, an eNB coordination protocol collects the requests and the scheduler dynamically assigns uplink transmission bitrates to the eNBs.

to and from the HeNB. In contrast, our proposed Sm-GW terminates all the control and data connections from HeNBs.

B. 3GPP Network Sharing

Network sharing was introduced by 3GPP in Technical Specification TS 23.951 [39] with the main motivation to share expensive radio spectrum resources among multiple operators. For instance, an operator without available spectrum in a particular geographic area can offer cellular services in the area through sharing the spectrum of another operator. In addition to spectrum sharing, 3GPP specifies core network sharing among multiple operators through a gateway core network (GWCN) configuration [39]. GWCN configurations are statically pre-configured at deployment for fixed pre-planned core network sharing. Thus, GWCN sharing can achieve only limited statistical multiplexing gain as the sharing is based on the pre-configured QoS requirements of the eNB interface connections and not on the varying eNB traffic demands. Also, the GWCN configuration lacks a central entity for optimization of the resource allocations with global knowledge of the eNB traffic demands. In contrast, our Sm-GW framework includes a central SDN orchestrator for optimized allocations of backhaul transmission resources according to the varying eNB traffic demands (see Section V).

III. PROPOSED SMART GATEWAY (SM-GW)

In this section we introduce the proposed Smart Gateway (Sm-GW) network architecture for existing LTE deployments. We describe the fundamental protocol mechanisms and interfaces that integrate the proposed Sm-GW into the conventional LTE protocols.

A. LTE Protocol Modifications

Fig. 4 illustrates the proposed protocol mechanisms between a set of N_s eNBs and a given Sm-GW s.

1) eNB: At the eNB, we newly introduce the eNB-to-Sm-GW reporting protocol, which operates on top of the GPRS tunneling protocol (GTP) [40] and stream control transmission protocol (SCTP). The reporting protocol (*i*) evaluates the required uplink transmission bitrate, and (*ii*) sends the bitrate request messages to the Sm-GW. The reporting protocol formulates the operator specific uplink transmission bitrate requests based on the requests of the UEs that are connected via the eNB to multiple operators $o,\ o = 1, 2, \ldots, O$.

The eNB wireless resource scheduler is responsible for the sharing of wireless resources between the eNB and the UEs. The eNB wireless resource scheduler ensures that only the resources available at the eNB are granted to the UEs. UEs periodically send buffer status reports (BSRs) to the eNB which they are connected to. Therefore, the eNB-to-Sm-GW reporting protocol can estimate the UE traffic requirements by interacting with the wireless resource scheduler.

2) Smart Gateway (Sm-GW): The protocol stack at the Sm-GW is similar to the HeNB-GW protocol stack. However, in the Sm-GW, an additional eNB coordination protocol, a scheduler for the dynamic resource allocation, and SDN capabilities are introduced.

The eNB coordination protocol collects request messages from eNBs. The eNB uplink transmission grants are sized based on the eNB requests and the available Sm-GW resources according to the Sm-GW scheduling described in Section IV. The eNB coordination protocol sends grant messages to all eNBs within a reasonable processing delay.

S1 based handovers for the downlink transmissions are typically anchored at the S-GW. (For the uplink transmissions, an anchoring, or buffering of packets, at a network entity, e.g., eNBs or S-GW, is not required.) We emphasize that the Sm-GW will be transparent to all the downlink packets from the S-GW and hence not be limited by the network protocol scheduler. This ensures that the S1 based handover mechanisms at the S-GW and eNBs continue to function normally.

3) SDN Operations of Sm-GW:

a) SDN infrastructure: The Sm-GW SDN capabilities can be provided by an OpenFlow (OF) agent and/or a configuration manager at each Sm-GW, as illustrated in Fig. 5. OpenFlow is a popular protocol for the southbound interface (SBI) and can be employed on the SBI between the Sm-GW SDN controller and Sm-GW. The OpenFlow agent supports OpenFlow SDN functionalities at the Sm-GW, making the Sm-GW configurable through the OpenFlow protocol.

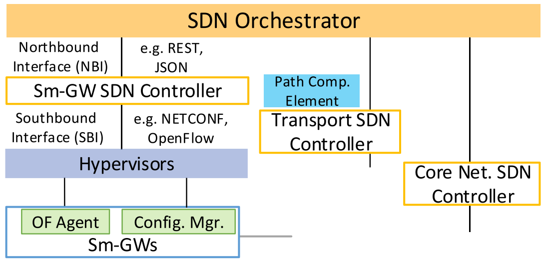


Fig. 5. Illustration of Sm-GW embedding in the SDN ecosystem: A set of Sm-GWs is controlled by an Sm-GW SDN controller. An SDN orchestrator coordinates the resource allocations and network management across the SDN controllers for the Sm-GWs, the transport network between Sm-GWs and operator gateways, and the operator core networks.

The Sm-GW configuration manager can be controlled by the Sm-GW SDN controller, e.g., through the NETCONF (or OpenFlow) SBI, to dynamically reconfigure the Sm-GW.

The Sm-GW SDN controller configures the Sm-GWs to enable the internal LTE X2 tunnel interfaces among all connected small cell eNBs, as elaborated in Section III-A4. Also, the Sm-GW SDN controller manages the external LTE X2 and S1 interfaces at the Sm-GW through tunnel establishments to the external LTE network core entities, i.e., MMEs and S/P-GWs.

Whereas the conventional LTE transport network between the eNBs and S/P-GWs is configured with static MPLS/IP paths [14], the flexible Sm-GW operation requires a flexible transport network, that is controlled by a transport SDN controller, as illustrated in Fig. 5. The flexible transport network can, for instance, be implemented through a Software Defined Elastic Optical Network (SD-EON) [41], [42].

b) Sm-GW virtualization: The SM-GW can support a variety of virtualization strategies, e.g., to provide independent virtual networks for different operators. One example virtualization strategy could let the Sm-GWs abstract the connected eNBs. Sm-GWs could then be abstracted by a hypervisor [43]–[46] that intercepts the SBI, as illustrated in Fig. 5. Based on the operator configurations that are sent via the SDN orchestrator to the Sm-GW SDN controller, resources at the Sm-GWs and the small cell eNBs (which are privately owned by an organization [6]) can be sliced to form operator-specific virtual networks of Sm-GWs and eNBs. The configuration manger at each Sm-GW can allocate resources to each of these virtual networks.

From the operator perspective, the Sm-GW virtualization essentially allows multiple operators to share the physical small cell infrastructure of Sm-GWs and eNBs. Thus, simultaneous services can be enabled to large UE populations that belong to multiple operators, i.e., that have contracts with multiple operators, while using the same physical small cell infrastructure. Existing conventional cellular deployment structures do not support the infrastructure sharing among multiple operators.

c) SDN orchestration: The SDN orchestrator coordinates the network management across multiple domains of Sm-GWs (whereby each Sm-GW domain is controlled by its own Sm-GW SDN controller), transport networks, and core networks. The SDN orchestrator implements the multi-operator management introduced in Section V and configures the Sm-GWs and transport networks based on global multi-operator network optimizations. For example, the SDN orchestrator communicates with the path computation element (PCE) SDN application on the transport SDN controller. The PCE dynamically evaluates the label switched paths, such as MPLS/IP paths, so as to flexibly enable and reconfigure the transport network [41], [42].

4) LTE X2 Interfaces of eNBs With Sm-GW: X2 interfaces enable critical functionalities in LTE small cells, such as X2-based handover as well as interference coordination and mitigation. Typically, each eNB connected to a given Sm-GW pertaining to an operator shares the same MME; thus, each eNB needs an X2 interface to all other eNBs within the same MME coverage area, the so-called tracking area. Hence, eNBs connected to an Sm-GW must be interconnected with X2 interfaces.

a) To external macro-eNBs: X2 traffic flows destined to eNBs located outside the scope of an Sm-GW (typically to a macro-eNB) are not limited by the scheduler at the Sm-GW. X2 packets flow out of the Sm-GW into the backhaul (i.e., to an S-GW) as they originate at the eNBs. The Sm-GW appears as an external router (or gateway) to the X2 external interfaces.

b) To internal small-eNBs: The Sm-GW appears as a simple bridge or a router to the internal X2 interfaces, routing the internal X2 packets within. Therefore, the scheduler at the Sm-GW does not limit any X2 packets. For small cell deployments, an eNB can have multiple neighboring eNBs in the tracking area; these neighboring eNBs need to be interconnected with each other with X2 connections. On the order of $O(N(N-1))$ dedicated links would be required to interconnect the X2 interfaces of N eNBs in the tracking area in a full mesh topology. In contrast, a star topology with the Sm-GW at the center requires only $O(N)$ connections to connect each eNB to the Sm-GW. In summary, in our Sm-GW architecture, the Sm-GW manages the X2 interfaces of all the internal small cell eNBs, thus eliminating the distributed management of X2 interfaces at each eNB.

5) Authentication of Sm-GW With EPC Core: Typically HeNBs use IPSec tunneling for their security and encryption, which creates overhead. If Sm-GWs are authenticated, the HeNBs would no longer need IPsec tunneling. Specifically, upon boot-up, the Sm-GW is authenticated with an LTE Evolved Packet Core (EPC) so as to eliminate the need for a network security-gateway (Se-GW) function or IPsec tunneling between the eNBs and the P-GWs. Critical cellular network functions, such as security, authentication, and reliability, require additional effort to be enabled in WiFi networks. WiFi Passpoint [47] (Hotspot 2.0) aims at providing an experience similar to cellular connectivity in WiFi networks by providing the cellular authentication mechanisms. With the authentication of Sm-GWs, the simplicity of WiFi networks can be achieved by the small cell cellular networks.

B. Downlink vs. Uplink Communication

1) Downlink Packets at the Sm-GW: Traffic flows in the conventional downlink path from an S/P-GW to an eNB are

TABLE I
SUMMARY OF NOTATION OF SM-GW NETWORK MANAGEMENT

Sm-GW Sched. Framework (Sm-GW ↔ eNBs), Sec. IV	
N_s	Number of small cell eNBs at Sm-GW s
G_{so}	Available uplink transm. bitrate [bit/s] from Sm-GW s to operator o
W	Duration [s] of scheduling cycle
Γ_{so}	$= G_{so}W/N_s$, Max. eNB uplink transm. data amount [bit] per cycle with equal sharing
ρ_{son}	Data amount [bit] that eNB n at Sm-GW s wants to transmit to operator o in a cycle, i.e., request by eNB n
γ_{son}	Data amount [bit] that eNB n at Sm-GW s is allowed to transmit to operator o in a cycle, i.e., grant by Sm-GW s
SDN Based Multi-Operator Managm. Framework, Sec. V (Sm-GWs ↔ Operator Gateways)	
o	Index of operators, $o = 1, 2, \dots, O$
s	Index of Sm-GWs, $s = 1, 2, \dots, S$
R_{so}	Smoothed uplink transmission bitrate [bit/s] request from Sm-GW s to operator o
K_o	Max. available uplink transm. bitrate through operator o
G_{so}	Granted uplink transm. bitrate from Sm-GW s to operator o
X_{so}	Actual uplink traffic bitrate from Sm-GW s to operator o

typically sent at rates that do not exceed the wireless transmission rates from the eNB to the UE devices. Thus, as long as the link rates from the S/P-GW to the inserted Sm-GW and from the Sm-GW to the eNB are at least as high as the conventional S/P-GW to eNB links, the Sm-GW can be transparent to the downlink packets from the S/P-GW.

2) Uplink Packets at Sm-GW: In contrast to the downlink data traffic, the uplink data traffic from the eNBs to an Sm-GW needs to be regulated as the traffic flows from all the eNBs terminating at the Sm-GW can overwhelm the outgoing link towards the operator S-GW. Enforcing QoS strategies and fairness among eNBs requires scheduling of the uplink packet traffic arriving from the eNBs at an Sm-GW. Therefore, our focus is on frameworks for the uplink transmission scheduling of the communication (*i*) from eNBs to an Sm-GW (Section IV), and (*ii*) from Sm-GWs to S-GWs (Section V).

IV. PROPOSED SM-GW SCHEDULING FRAMEWORK

A. Purpose

The main purpose of the Sm-GW scheduling framework is to maximize the utilization of the network resources, and to ensure fair uplink transmission service for all eNBs connected to an Sm-GW. Without scheduling, highly loaded eNBs can impair the service for lightly loaded eNBs connected to the same Sm-GW. When many eNBs are flexibly connected to an Sm-GW, traffic bursts from heavily loaded eNBs can overwhelm the queue of an Sm-GW, resulting in excessive packet drops and high delays, even for lightly loaded eNBs. On the other hand, with scheduling, a large number of eNBs can be flexibly connected to the Sm-GW while ensuring prescribed QoS and fairness levels. Each eNB can possibly have a different service level agreement. The Sm-GW allows for the flexible deployment of a wide variety of scheduling algorithms. We outline two classes of Sm-GW scheduling algorithms, and illustrate an elementary algorithm for each class. Table I summarizes the main Sm-GW framework notations.

B. Configuration Adaptive Scheduling

Configuration adaptive scheduling adapts the scheduling, i.e., the allocation of uplink transmission bitrates, according to the number of eNBs connected to a given Sm-GW. The Sm-GW tracks the number of connected eNBs and sends a configuration message to all eNBs in the event of a change in connectivity at the Sm-GW, i.e., addition of new eNB or disconnection of existing eNB. More specifically, consider N_s eNBs at a given Sm-GW s that has been allocated the uplink transmission bitrate G_{so} [bit/s] toward a given operator o (through the coordination techniques in Section V).

An elementary equal share scheduling shares the available uplink transmission bitrate at the Sm-GW toward a given operator o equally among all eNBs connected to the Sm-GW. Each eNB n, $n = 1, 2, \dots, N_s$, can then transmit at most $\Gamma_{so} = G_{so}W/N_s$ [Byte] of traffic during a cycle of duration W [seconds]. The traffic amount limit Γ_{so} and cycle duration W are sent to the eNBs as a part of the initial configuration message. Each eNB schedules the uplink transmissions such that no more than Γ_{so} [Byte] of traffic are send in a cycle of duration W [seconds].

The simple equal share scheduler can flexibly accommodate large numbers N_s of eNBs. However, the equal bandwidth assignments by the elementary equal share scheduler to the eNBs under-utilize the network resources when some eNBs have very little traffic while other eNBs have high traffic loads.

C. Traffic Adaptive Scheduling

With traffic adaptive scheduling, the Sm-GW collects uplink transmission requests from the eNBs. The Sm-GW then adaptively allocates portions of the uplink transmission bitrate G_{so} to the individual eNBs according to their requests. Traffic adaptive scheduling operates with a request-allocate-transmit cycle of duration W [seconds] illustrated in Fig. 6. At the start of the cycle, each eNB n, $n = 1, 2, \dots, N_s$, sends an uplink transmission bitrate request to Sm-GW s. We let ρ_{son} denote the amount of traffic [in Byte] that eNB n wants to transmit to operator o over the next cycle of duration W. Once all requests have been received, i.e., following the principles of the offline scheduling framework [48], portions of G_{so} can be allocated to the eNBs according to some scheduling policy.

An elementary excess share scheduling policy [49] allocates the eNB grants as follows. Lightly loaded eNBs with $\rho_{son} < \Gamma_{so}$ are granted their full request, i.e., receive the grant size $\gamma_{son} = \rho_{son}$, while their unused (excess) portion of the equal share allocation is accumulated in an excess pool:

$$\xi = \sum_{\forall \rho_{son} \le \Gamma_{so}} \Gamma_{so} - \rho_{son}. \tag{1}$$

Following the principles of controlled equitable excess allocation [49], highly loaded eNBs are allocated an equal share of the excess up to their request. That is, with $|\mathcal{H}|$ highly loaded eNBs, the grants are

$$\gamma_{son} = \min\left(\rho_{son},\ \Gamma_{so} + \frac{\xi}{|\mathcal{H}|}\right). \tag{2}$$

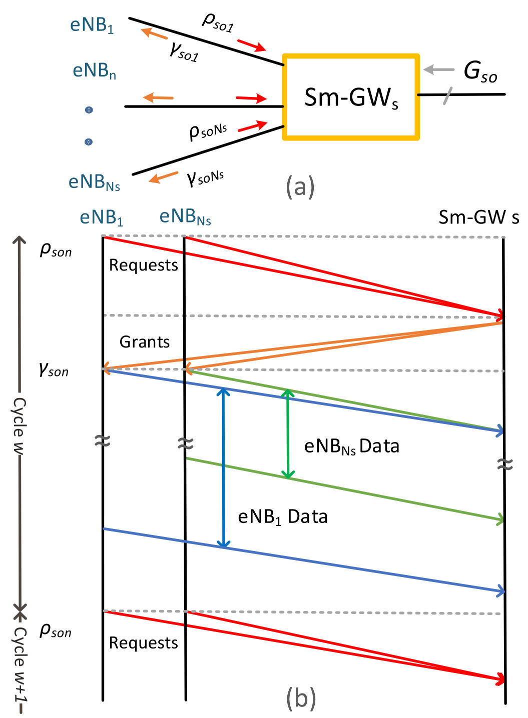


Fig. 6. Illustration of traffic adaptive Smart Gateway (Sm-GW) scheduling: (a) Sm-GW s receives uplink transmission bitrate allocation G_{so} from operator o. Based on uplink transmission bitrate requests from individual eNBs n, $n = 1, \ldots, N_s$, Sm-GW s sends uplink transmission grants to the eNBs. (b) The eNB requests and Sm-GW grants are followed by the eNB uplink data transmissions (in parallel) in a fixed-period cycle.

D. Scheduling Fairness

Within the context of our proposed Sm-GW scheduling framework, fairness is the measure of network accessibility of all N_s eNBs connected to an Sm-GW s based on individual eNB uplink throughput level requirements. We denote T_{son} for the long-run average throughput level [bit/s] of uplink traffic generated at eNB n, $n = 1, 2, \ldots, N_s$, at Sm-GW s for operator o. The throughput level T_{son} can for instance be obtained through smoothing of the requests ρ_{son} over successive cycles w. In order to avoid clutter, we omit the subscripts s and o in the remainder of this fairness evaluation. We define the following fair target throughput levels Ω_n [bit/s]: Lightly loaded eNBs $l \in \mathcal{L}$ with throughput levels $T_l < \Gamma/W$, should be able to transmit their full traffic load, i.e., $\Omega_l = T_l$. Next, consider highly loaded eNBs $h \in \mathcal{H}$ with throughput levels $T_h > \Gamma/W$. If the total throughput requirement of all eNBs $\sum_{l\in\mathcal{L}} T_l + \sum_{h\in\mathcal{H}} T_h$ is less than or equal to the uplink transmission bitrate G, then the highly loaded eNBs should be able to transmit their full traffic load, i.e., $\Omega_h = T_h$. On the other hand, if the total traffic load exceeds the uplink transmission bitrate, i.e., if $\sum_{l\in\mathcal{L}} T_l + \sum_{h\in\mathcal{H}} T_h > G$, then the highly loaded eNBs should be able to transmit traffic up to an equitable share of the uplink transmission bitrate not used by the lightly loaded eNBs. Thus, overall: $\Omega_h = \min\{T_h, \ (G - \sum_{l\in\mathcal{L}} T_l)/|\mathcal{H}|\}$. We define the normalized distance $\mathcal{E}_n$ of the actually achieved (observed) throughput τ_n and the target throughput Ω_n, i.e., $\mathcal{E}_n = \tau_n - \Omega_n$.

Based on the preceding target throughput definitions, we obtain the normalized distance throughput fairness index [50]

$$\mathcal{F}_T = \frac{\sqrt{\sum_{n=1}^{N} \mathcal{E}_n^2}}{\sqrt{\sum_{n=1}^{N} \Omega_n^2}}, \tag{3}$$

whereby $\mathcal{F}_T$ close to zero indicates fair Sm-GW scheduling.

E. Sm-GW Scheduling Overhead

In configuration adaptive Sm-GW scheduling, a reconfiguration event, i.e., an eNB connect or disconnect event, triggers the re-evaluation of the grant size limit Γ_{so}, see Section IV-B. The new Γ_{so} value is sent to all eNBs. Since reconfiguration events occur typically only rarely, e.g., on the time scale of minutes or hours, the overhead for configuration adaptive scheduling is negligible.

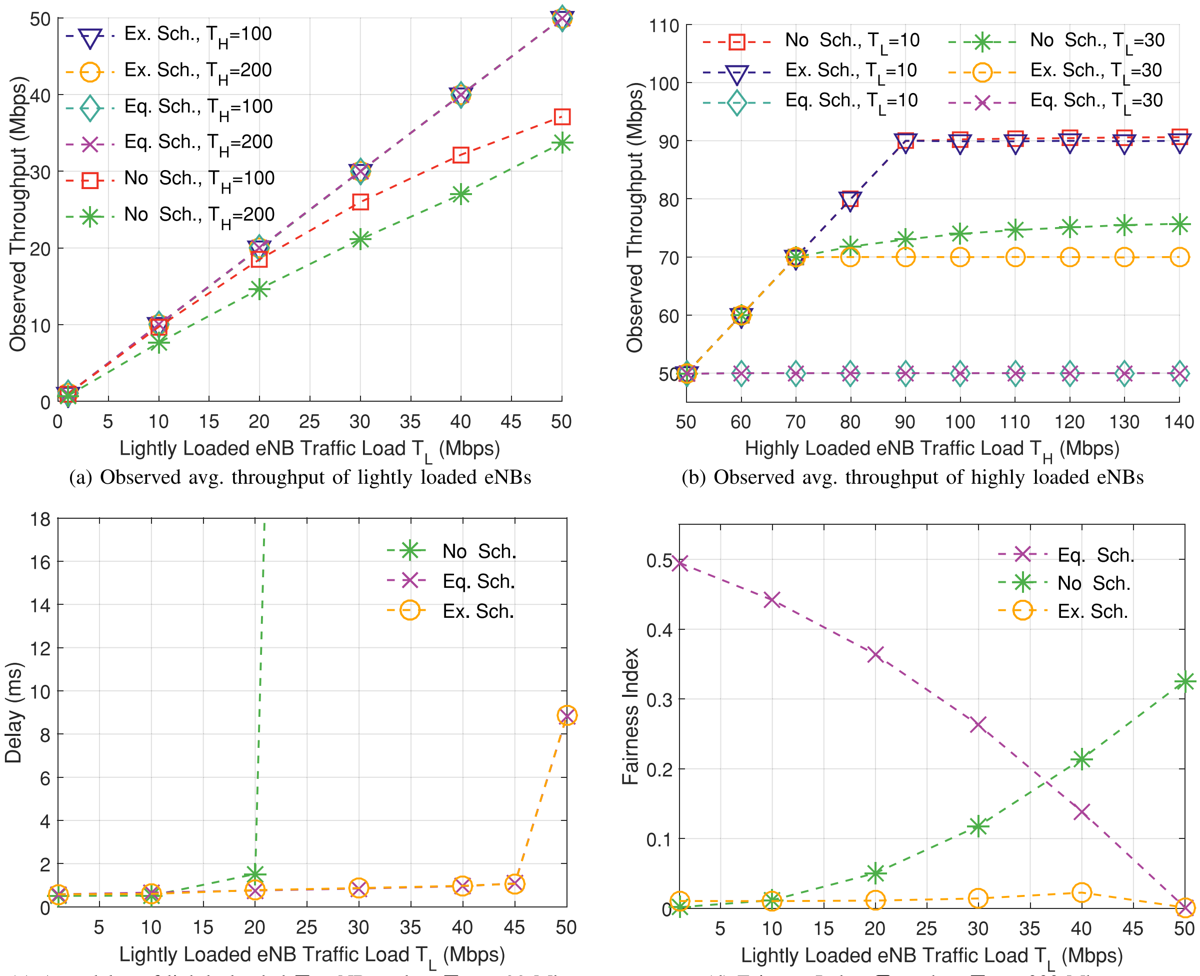
Traffic adaptive Sm-GW scheduling requires each eNB n to send a request every cycle of duration W seconds. Upon reception of the requests from all N_s eNBs, the Sm-GW evaluates and sends the grants to the respective eNBs, as illustrated in Fig. 6(a). The requests and grants can typically be sent simultaneously, i.e., in parallel, over the individual eNB-to-Sm-GW links. Thus, within a cycle duration W, the overhead amounts to the transmission delays of the request and grant messages, the maximum round-trip propagation delay between eNBs and Sm-GW, and the schedule processing delay at the Sm-GW. For typical parameter settings, such as 70 Byte messages transmitted at 1 Gbps, up to 500 m eNB-to-Sm-GW propagation distance, $W = 1$ ms cycle duration, and schedule processing delay on the order of microseconds, the overhead is less than half a percent.

F. Evaluation of Sm-GW Scheduling

1) Simulation Setup: We evaluate the performance of Sm-GW scheduling with the discrete event simulator OMNET++. We consider a given Sm-GW s with an uplink transmission bitrate to a given operator o of $G_{so} = 1$ Gbps. We omit the subscripts s and o in the remainder of this evaluation section to avoid notational clutter. The LTE access network typically requires the packet delay to be less than 50 ms [51]. Therefore, we set the Sm-GW queue size to 20 MBytes, which is equivalent to a maximum queuing delay of 20 ms over the $G = 1$ Gbps link. Without any specific scheduling, the Sm-GW operates in first-come-first-served mode with taildrop.

We simulate the typical bursty eNB traffic generation pattern, with two eNB traffic rate states: low and heavy. The sojourn time in a given traffic rate state is randomly drawn from a uniform distribution over 1 ms to 4 ms. At the end of the sojourn time, a switch to another state occurs with a probability of 70 % in the low traffic state and 30 % in the heavy traffic state. The traffic bitrate ratio between the heavy and low traffic states is 4:1. Within a given traffic rate state, data packets are randomly generated according to independent Poisson processes.

We consider $|\mathcal{L}| = 10$ lightly loaded eNBs and $|\mathcal{H}| = 10$ highly loaded eNBs connected to the considered Sm-GW. Each eNB, irrespective of whether it is lightly or highly loaded,



(a) Observed avg. throughput of lightly loaded eNBs

(b) Observed avg. throughput of highly loaded eNBs

(c) Avg. delay of lightly loaded T_L eNBs, when $T_H = 80$ Mbps

(d) Fairness Index $\mathcal{F}_T$, when $T_H = 200$ Mbps.

Fig. 7. Simulation results for Sm-GW scheduling: Actual (observed) long-run average throughput levels τ of lightly and highly loaded eNBs, packet delay of lightly loaded eNBs, and fairness index $\mathcal{F}_T$ as a function of long-run eNB traffic load (requirement). Fixed parameters: $G = 1$ Gbps Sm-GW uplink transmission bitrate, $|\mathcal{L}| = 10$ lightly loaded and $|\mathcal{H}| = 10$ highly loaded eNBs.

generates traffic according to the two traffic rate state (low and heavy) model. The low and heavy traffic rates are set such that the long-run average generated traffic rate corresponds to a prescribed required throughput (load) level $T_L < G/N = 50$ Mbps for a lightly loaded eNB and a prescribed required throughput (load) level $T_H > G/N$ for a highly loaded eNB. For all simulations, the 95 % confidence intervals are less than 5 % of the corresponding sample mean.

2) Simulation Results:

a) Without Sm-GW scheduling: In Fig. 7, we show representative evaluation results comparing configuration adaptive equal-share Sm-GW scheduling and traffic adaptive excess-share Sm-GW scheduling with the conventional backhaul without Sm-GW scheduling. Figs. 7(a) and (b) show the actual (achieved, observed) throughput τ of lightly loaded and highly loaded eNBs, respectively, as a function of the generated lightly loaded (T_L) and highly loaded (T_H) throughput levels. We observe from Figs. 7(a) that without scheduling, the lightly loaded eNB suffer reductions in the achieved throughput, that are especially pronounced (over 30 %) for the high $T_H = 200$ Mbps load of the highly loaded eNBs. At the same time, we observe from Figure 7(b) that without scheduling, the highly loaded eNBs achieve more than their fair throughput share. For instance, for the highly loaded eNB throughput requirement (load) $T_H = 140$ Mbps, and $T_L = 30$ Mbps, the observed throughout of the highly loaded eNBs is $\tau_H = 76$ Mbps, which is significantly higher than the fair share of $(G - |\mathcal{L}|T_L)/|\mathcal{H}| = 70$ Mbps. The unfairness arising without scheduling is further illustrated in Fig. 7(c), where we observe a sharp delay increase at $T_L = 20$ Mbps, when the total traffic load $|\mathcal{L}|T_L + |\mathcal{H}|T_H$ approaches the uplink transmission bitrate G. Moreover, from Fig. 7(d), we observe an increasing fairness index $\mathcal{F}_T$ as the lightly loaded eNBs generate more traffic, i.e., as T_L increases. That is, as the lightly loaded eNBs try to transmit more traffic, their achieved throughput falls more and more below their fair share [see growing divergence between the no scheduling curves and straight lines for scheduling in Fig. 7(a)], leading to increasingly unfair treatment of the lightly loaded eNBs.

b) Equal-share Sm-GW scheduling: We observe from Fig. 7(a) and (c) that lightly loaded eNBs benefit from equal-share scheduling in that they get the full share of their fair target throughput and experience low delay. However, we observe from Fig. 7(b) that highly loaded eNBs achieve only a throughput of $G/(|\mathcal{L}| + |\mathcal{H}|) = 50$ Mbps as equal-share Sm-GW scheduling assigns a configuration adaptive allocation of equal shares of the limited uplink transmission bitrate G to all eNBs irrespective of their traffic generation rates. Correspondingly, we observe from Fig. 7(d), a high fairness index $\mathcal{F}_T$ for low traffic loads of the lightly loaded eNBs, as the highly loaded eNBs receive only unfairly small shares of the uplink transmission bitrate G.

c) Excess-share Sm-GW scheduling: We observe from Fig. 7(a) and (b) that with excess-share Sm-GW scheduling, both lightly loaded eNBs and highly loaded eNBs achieve their fair target throughput. We further observe from Figs. 7(c) and (d) that excess-share Sm-GW scheduling gives also favorable delay and fairness index performance.

d) Summary: We conclude that scheduling of the Sm-GW uplink transmission bitrate G is necessary to prevent backhaul bandwidth starvation of lightly loaded eNBs due to the overwhelming traffic rates of highly loaded eNBs. On the other hand, simple configuration adaptive allocation of equal uplink transmission bitrate shares to each eNB wastes bandwidth. Flexible traffic adaptive scheduling according to the traffic loads of the eNBs, e.g., through excess-share scheduling, can ensure fairness while efficiently utilizing the uplink transmission bitrate.

V. SDN Based Multi-Operator Management

A. Overview

In this section we introduce a novel SDN based network management framework for flexible sharing of the backhaul resources of multiple operators. In particular, the framework introduced in Sections V-B–V-E allows a set of Sm-GWs to flexibly share the uplink transmission bitrate of a given single operator. The inter-operator sharing introduced in Section V-F allows the Sm-GWs to flexibly share the uplink transmission bitrates of multiple operators. Our proposed multi-operator management framework accommodates dynamic changes of the traffic requirements of the small cells, such as changes of the generated uplink traffic bitrates, as well as dynamic changes of the operator characteristics, such as changes of the available uplink traffic bitrates. In the proposed multi-operator management framework, an SDN orchestrator dynamically configures the Sm-GWs and the transport network connecting the Sm-GWs to the operator gateways to flexibly adapt to changes in small cell traffic loads and the operator characteristics.

B. Request and Allocation Procedures

In a small cell deployment environment, such as a large organization, multiple Sm-GWs can serve multiple buildings. For example, in a university setting, a library can be equipped with an Sm-GW and the administration building can be equipped with another Sm-GW. The throughput requirements and priorities of these buildings typically vary widely over time. For instance, the administration building experiences a large visitor influx during graduation and student admission periods, while many students visit the library during exam week. Moreover, services from multiple operators may need to be shared among the buildings in a given organization, i.e., among multiple Sm-GWs. Hence, there is a need for highly flexible traffic management within the large organization based on time-varying priorities and throughput requirements.

Suppose, the Sm-GWs s, $s = 1, 2, \ldots, S$, and operators o, $o = 1, 2, \ldots, O$, are interconnected in a full mesh transport network, as illustrated in Fig. 8. As described in Section IV-C, with traffic adaptive Sm-GW scheduling, each eNB n sends its operator o specific uplink transmission bitrate request to Sm-GW s in every cycle. The requested uplink transmission data amounts ρ_{son} will typically vary over time scales that are long enough to reasonably permit adaptations of the Sm-GW configurations. For instance, the requests will typically change on the time scales of several seconds or minutes, or possibly even longer, such as the seasonal variations in the visitor volume to university buildings. In order to obtain the variational characteristics of the eNB requirements, the operator specific requests at the Sm-GWs can be aggregated over the eNBs and appropriately smoothed, i.e., averaged over time, to obtain an aggregate smoothed uplink transmission bitrate request R_{so} [bit/s] from Sm-GW s to operator o.

Ideally, the backhaul network should adapt to varying requirements at the Sm-GWs to maximize the network utilization. We exploit the centralized control property of SDN to adaptively configure the network for variable requirements. More specifically, the SDN orchestrator in Fig. 8 optimizes the allocations G_{so} of operator o uplink transmission bitrate [bit/s] to the individual Sm-GWs s. The SDN orchestrator thus ensures that the grants to the eNBs are globally maximized (subject to the operators' constraints and requirements). When the optimized allocations G_{so} are used at the Sm-GW schedulers (Section IV), the maximum allowed traffic flow is sent from the Sm-GWs to each operator core.

C. Optimization Decision Variables and Constraints for Multi-Operator Management With Sm-GWs

In this section, we define a general optimization model for the multi-operator management framework. Specifically, we define the constraints and decision variables for optimizing the multi-operator management. The defined decision variables and constraints are employed for the operation of the SDN orchestrator, as detailed in Section V-D. The SDN orchestrator can employ arbitrary objective functions and constraint allocation strategies for the optimization, as illustrated for an elementary example in Section V-E.

1) Constraints: Requests for the uplink transmission of ρ_{son} [bits] from eNBs n, $n = 1, 2, \ldots, N_s$, arrive at Sm-GW s, $s = 1, 2, \ldots, S$, every cycle of duration W seconds, i.e., on the order of milliseconds, requesting uplink transmission bitrates from operator o, $o = 1, 2, \ldots, O$. The Sm-GW aggregates the requests over the eNBs n and smoothes the aggregated requests to obtain the smoothed aggregated requests R_{so}. Denoting w for the cycle index, an elementary weighted sampling smoothing computes

$$R_{so}(w) = \alpha \left(\frac{1}{N_s} \sum_{n=1}^{N_s} \frac{\rho_{son}(w)}{W} \right) + (1 - \alpha)\, R_{so}(w - 1), \qquad (4)$$

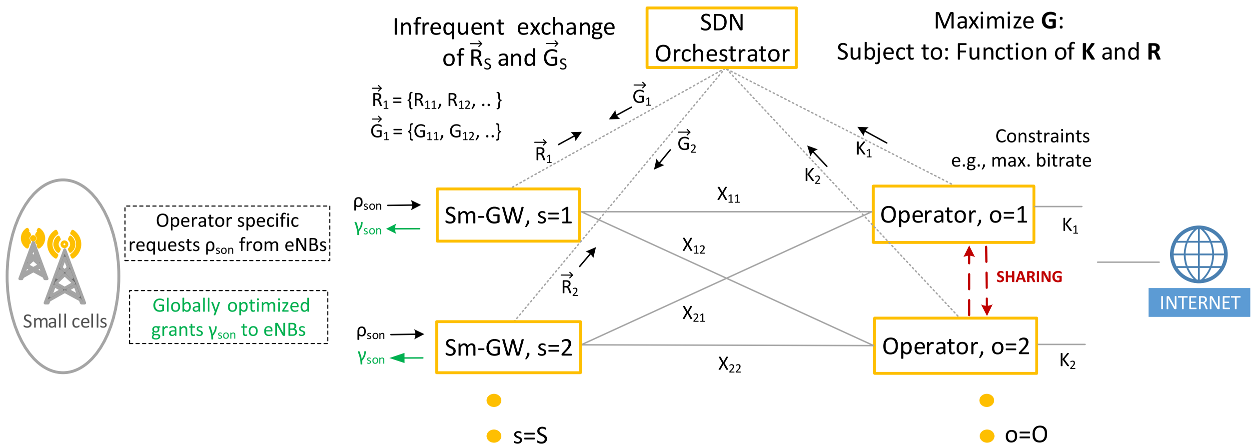


Fig. 8. Illustration of SDN based multiple-operator management serving multiple Smart Gateways (Sm-GWs): Two operators $o = 1$ and $o = 2$ with uplink transmission bitrate constraints K_1 and K_2 provide services through two Sm-GWs $s = 1$ and $s = 2$ to a small cell eNB infrastructure. Uplink transmission requests ρ_{son} from individual eNBs n are aggregated and averaged over time to obtain Sm-GW s uplink transmission bitrate requests R_{so} for each operator o. The SDN orchestrator optimizes the grants G_{so} of the limited operator uplink transmission bitrates K_o to the individual Sm-GWs s. The Sm-GW s in turn uses the granted uplink transmission bit rates G_{so} for the scheduling of the individual eNB uplink transmissions (see Section IV). The SDN orchestration maximizes the Sm-GW s to operator o traffic flow rates X_{so}.

where α denotes the weight for the most recent request sample. A wide variety of other smoothing mechanism can be employed and optimized according to the specific deployment settings. The smoothed requests R_{so} are periodically (with a period typically much longer than the eNB reporting window) sent to the SDN orchestrator. In particular, each Sm-GW s, sends a vector of smoothed requests $\overrightarrow{R_s} = [R_{s1}\ R_{s2}\ \cdots\ R_{sO}]$ containing the aggregated and smoothed requests for each operator o to the SDN orchestrator. The SDN orchestrator combines the request vectors $\overrightarrow{R_s}$ to form the request matrix

$$\mathbf{R} = [R_{so}],\ \ s = 1, 2, \ldots, S;\ \ o = 1, 2, \ldots, O. \tag{5}$$

Each operator $o,\ \ o = 1, 2, \ldots, O$, can enforce a set of constraints $K_{oc},\ \ c = 1, 2, \ldots, C$, represented by a constraint vector $\overrightarrow{K_o} = [K_{o1}\ K_{o2}\ \cdots\ K_{oC}]$ that is sent to the SDN orchestrator. Each constraint c may be associated with a particular specification from operator o, e.g., for traffic shaping of the flows or for the aggregate maximum bitrate. In order to avoid clutter and not to obscure the main ideas of our overall multi-operator management framework, we consider in this study a single constraint for each operator o. That is, in place of the constraint vector $\overrightarrow{K_o}$ we consider a single (scalar) constraint K_o. The SDN orchestrator combines the scalar constraints from the various operators o to form the constraint vector

$$\mathbf{K} = [K_1\ K_2\ \cdots\ K_O]. \tag{6}$$

2) Decision Variables: The Sm-GW s scheduler uses the operator o specific grant size limits Γ_{so} to schedule/assign uplink transmission grants to eNBs (see Sections IV-B and IV-C). By controlling the variable Γ_{so} specific to operator o we can control the flow of traffic outward from the Sm-GW, i.e., towards the respective operator o. The long-term average traffic flow rates X_{so} [bit/s] from the Sm-GWs $s,\ \ s = 1, 2, \ldots, S$, to the operators $o,\ \ o = 1, 2, \ldots, O$, can be expressed as matrix

$$\mathbf{X} = [X_{so}],\ \ s = 1, 2, \ldots, S; o = 1, 2, \ldots, O. \tag{7}$$

The operator o specific uplink transmission bitrates G_{so} granted to the Sm-GWs are evaluated at the SDN orchestrator, based on the request matrix $\mathbf{R}$ and the constraint vector $\mathbf{K}$. The orchestrator responds to the request vector $\overrightarrow{R_s}$ from each Sm-GW s with a grant vector $\overrightarrow{G_s}$. At the SDN orchestrator, the grant vectors $\overrightarrow{G_s}$ can be combined to form the orchestrator grant matrix

$$\mathbf{G} = [G_{so}],\ \ s = 1, 2, \ldots, S; o = 1, 2, \ldots, O. \tag{8}$$

$\mathbf{G}$ is a positive (non-negative) matrix, since the matrix elements $G_{so},\ \ G_{so} \geq 0$, correspond to granted uplink transmission bitrates.

Our objective is to maximize the traffic flow rates X_{so} from the Sm-GWs s to the operators o subject to the operator constraints $\mathbf{K}$. In particular, the aggregated traffic sent from the Sm-GWs $s,\ \ s = 1, 2, \ldots, S$, to the operator o core should satisfy the operator constraint K_o, i.e.,

$$\sum_{s=1}^{S} X_{so} \leq K_o, \quad \forall o,\ \ o = 1, 2, \ldots, O. \tag{9}$$

Using the grant vector $\overrightarrow{G_s}$ at Sm-GW s to assign, i.e., to schedule, uplink traffic grants to the eNBs (see Section IV) ensures that the traffic flow rates X_{so} from Sm-GW s to operator o are bounded by G_{so}, i.e.,

$$X_{so} \leq G_{so}, \quad \forall (s, o). \tag{10}$$

Thus, in order to ensure that the traffic flows X_{so} satisfy the operator constraints $\mathbf{K}$, the grants G_{so} must satisfy the operator constraints, i.e.,

$$\sum_{s=1}^{S} G_{so} \leq K_o, \quad \forall o,\ \ o = 1, 2, \ldots, O. \tag{11}$$

In order to maximize the traffic flows X_{so} to each operator o, the SDN orchestrator needs to grant each Sm-GW s the maximum permissible uplink transmission bitrate G_{so}.

Algorithm 1: SDN Orchestrator Procedure

1. Sm-GWs
(a) Evaluate aggregate smoothed requests R_{so} from eNB requests ρ_{son}, Eqn. (4)
(b) Periodically send request vector $\overrightarrow{R_s}$ to SDN orchestrator
if *Grant vector* $\overrightarrow{G_s}$ *is received* **then**
| Update Sm-GW (to eNBs) grant size limits Γ_{so}
end
2. Operators
(a) Send constraint K_o to SDN orchestrator
3. SDN Orchestrator
if *Request vector* $\overrightarrow{R_s}$ *is received* ***OR*** *constraint* K_o *is received* **then**
| Re-optimize orchestrator (to Sm-GW) grants **G**
| Send grant vector $\overrightarrow{G_s}$ to Sm-GW s
end

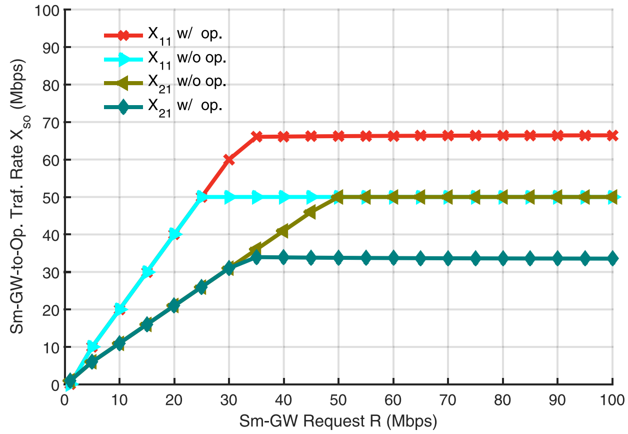


Fig. 9. Traffic rates X_{11} and X_{21} from Sm-GWs $s = 1$ and $s = 2$ to operator $o = 1$ as a function of Sm-GW request rate R for Sm-GW requests $\mathbf{R} = [R_{11}\ \ R_{12};\ \ R_{21}\ \ R_{22}] = [2R\ \ 50;\ \ R\ \ 50]$ Mbps without and with SDN orchestrator optimization (without inter-operator sharing); fixed operator uplink transmission bitrate constraints $\mathbf{K} = [K_1\ \ K_2] = [100\ \ 100]$ Mbps.

D. SDN Orchestrator Operation

The operational procedures for evaluating the SDN orchestrator grant matrix **G** (8) are executed in parallel in the Sm-GWs, operators, and the SDN orchestrator, as summarized in Algorithm 1. The Sm-GWs aggregate and smooth the eNB requests and periodically send the request vector $\overrightarrow{R_s}$ to the SDN orchestrator. The SDN orchestrator optimizes the grant matrix **G** upon the arrival of a new Sm-GW request vector $\overrightarrow{R_s}$ or a change in an operator constraint K_o. The orchestrator updates the Sm-GWs with the newly evaluated orchestrator grant vectors $\overrightarrow{G_s}$, which update their grant size limits Γ_{so}.

Our SDN based multi-operator management framework allows for a wide variety of resource (uplink transmission bitrate) allocations from the multiple operators to the Sm-GWs. In order to illustrate the introduced framework, we consider next an elementary specific optimization problem formulation with a linear objective function and a proportional constraint allocation strategy that allocates the uplink transmission bitrate constraints proportional to the requests. More complex objective functions and allocation strategies, e.g., objective functions that prioritize specific grants, are an interesting direction for future research. We note that this illustrative example does not exploit inter-operator sharing, which is examined in Section V-F.

E. Illustrative Optimization Example With Linear Objective Function and Request-Proportional Constraint Allocations

Since the grants G_{so} are non-negative, an elementary objective function can linearly sum the grants G_{so}, i.e., as $\sum_{s=1}^{S}\sum_{o=1}^{O} G_{so}$. For the constraint allocation, we consider the aggregate over all Sm-GWs s of the aggregated smoothed requests R_{so} for a specific operator o, i.e., we consider the unit norm of the request vector $\|\overrightarrow{R_o}\|_1 = \sum_{s=1}^{S} R_{so}$. If $\|\overrightarrow{R_o}\|_1$ is less than the operator constraint K_o, then the corresponding grants G_{so} are set to the requests, i.e., $G_{so} = R_{so}$. On the other hand, if $\|\overrightarrow{R_o}\|_1 > K_o$, then we proportionally assign the operator o backhaul bandwidth K_o, i.e., we assign the proportion $R_{so}/\|\overrightarrow{R_o}\|_1$ of the constraint K_o. Thus,

$$G_{so} = \min\left(R_{so},\ \frac{R_{so}}{\left\|\overrightarrow{R_o}\right\|_1} K_o\right). \tag{12}$$

The resulting elementary optimization problem can be summarized as:

$$\begin{aligned} &\text{Maximize} \quad \sum_{s=1}^{S}\sum_{o=1}^{O} G_{so} \\ &\text{Subject to:} \quad \forall s \in \{1, 2, \ldots, S\} \text{ and } \forall o \in \{1, 2, \ldots, O\}, \\ &\qquad -\ G_{so} \le 0, \\ &\qquad G_{so} \le R_{so}, \\ &\qquad G_{so} \le K_o \frac{R_{so}}{\left\|\overrightarrow{R_o}\right\|_1}. \end{aligned} \tag{13}$$

F. Inter-Operator Sharing

When the aggregate backhaul bandwidth $\|R_o\|_1$ requested from an operator o exceeds its constraint K_o, inter-operator sharing can be employed to route the additional traffic through the network managed by another operator. Our proposed Sm-GW multi-operator management provides a distinctive advantage in maintaining active connections with multiple operators to easily route the excess traffic to a sharing operator. We denote $o = m$ for the operator that accepts the sharing traffic from an other operator $o = e$ whose traffic constraint has been exceeded. In this study, we focus on one operator accepting sharing traffic and one operation with excess traffic. The extension to sets of multiple operators accepting sharing traffic and multiple operators with excess traffic is left for future research. An operator in sharing m should have low incoming traffic as compared to the contraint K_m in order to accept the traffic from the operator in excess e. Therefore, for the sharing

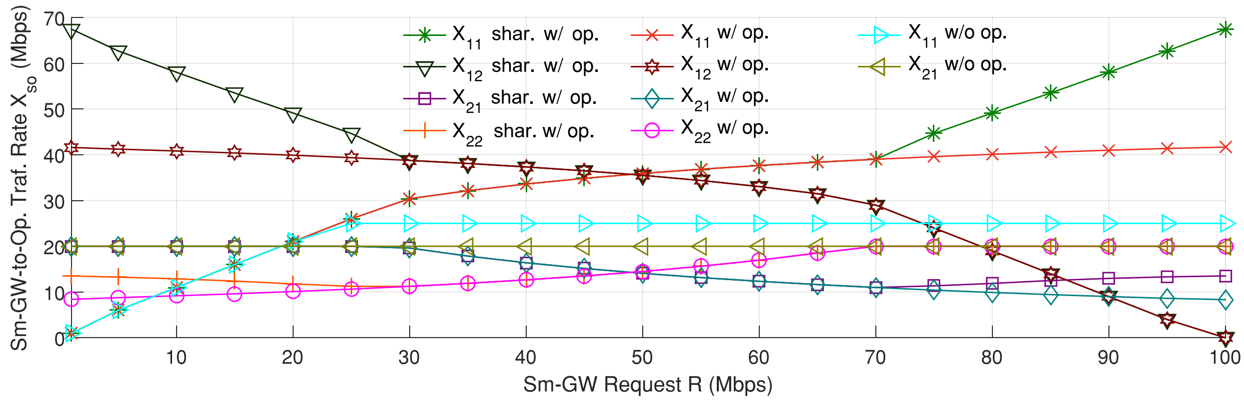


Fig. 10. Inter-operator sharing evaluation: Traffic rates X_{so} from Sm-GWs s, $s = 1, 2$, to operators o, $o = 1, 2$, as a function of Sm-GW request rate R for Sm-GW requests $\mathbf{R} = [R_{11} \;\; R_{12}; \;\; R_{21} \;\; R_{22}] = [R \;\; 100 - R; \;\; 20 \;\; 20]$ Mbps with and without SDN orchestrator optimization; fixed operator uplink transmission bitrate constraints $\mathbf{K} = [K_1 \;\; K_2] = [50 \;\; 50]$ Mbps.

($o = m$) and excess ($o = e$) operators the requests R_{so} need to satisfy,

$$\sum_{s=1}^{S} R_{sm} < K_m, \text{ and } \sum_{s=1}^{S} R_{se} > K_e. \qquad (14)$$

The traffic rate from excess operator e that can be carried by sharing operator m depends on the unutilized slack uplink transmission bitrate of operator m:

$$\zeta = K_m - \sum_{s=1}^{S} R_{sm}. \qquad (15)$$

If $\zeta > 0$, the last constraint in optimization problem (13) for the excess operator e is replaced by the constraint

$$G_{se} \leq (K_e + \zeta) \frac{R_{se}}{\left\| \overrightarrow{R_e} \right\|_1} \quad \forall s. \qquad (16)$$

G. Evaluation of Multi-Operator Management

In order to showcase the effectiveness of the SDN based multi-operator management framework, we conducted simulations for the elementary optimization with linear objective function and proportional constraint sharing (see Section V-E). We consider $S = 2$ Sm-GWs and $O = 2$ operators. As comparison benchmark, we consider a static equal allocation of operator uplink transmission bitrate K_o to the S Sm-GWs, i.e., each Sm-GW s, $s = 1, 2$, is allocated K_o/S of the operator o uplink transmission bitrate.

1) Without Inter-Operator Sharing: In Fig. 9 we plot the Sm-GW s to operator o traffic flow rates X_{so} resulting from the optimized SDN orchestrator grants G_{so} as a function of the uplink transmission bitrate requested by Sm-GWs $s = 1$ and $s = 2$ from operator $o = 1$. Specifically, Sm-GW $s = 1$ requests bitrate $R_{11} = 2R$ and Sm-GW $s = 2$ requests bitrate $R_{21} = R$ from operator $o = 1$. The bitrate requests from operator $o = 2$ are fixed at 50 Mbps. Each operator o, $o = 1, 2$, has uplink transmission bitrate constraint $K_o = 100$ Mbps.

We observe from Fig. 9 that for requests for operator $o = 1$ bitrate up to $R = 25$ Mbps, the traffic rates X_{11} and X_{21} are equal to the requests, irrespective of whether SDN orchestrated optimization is employed or not. In contrast, as the requested bitrate increases above $R = 25$ Mbps, i.e., the bitrate $R_{11} = 2R$ requested by Sm-GW $s = 1$ from operator $o = 1$ increases above $K_1/S = 50$ Mbps, the granted bitrate G_{11} with SDN orchestration and the corresponding traffic flow X_{11} continue to increase. On the other hand, the granted bitrate G_{11} and traffic flow X_{11} without SDN orchestration stay limited at the static equal share $X_{11} = G_{11} = K_1/S = 50$ Mbps.

As the requested bitrate R increases above 33.3 Mbps, i.e., a total of $3R = 100$ Mbps requested from operator $o = 1$, we observe from Fig. 9 that without orchestration, the traffic flow X_{21} from Sm-GW $s = 2$ to operator $o = 1$ grows to and then remains at the static equal share $K_1/S = 50$ Mbps. That is, the conventional static uplink transmission bitrate allocation results in unfair disproportional backhaul service. In contrast, our dynamic multi-operator management with SDN orchestrated optimization based on proportional sharing adaptively assigns the operator $o = 1$ bitrate to Sm-GWs $s = 1$ and $s = 2$ proportional to their requests.

2) With Inter-Operator Sharing: In Fig. 10, we plot the Sm-GW s to operator o traffic flow rates X_{so} as a function of the uplink transmission bitrate $R_{11} = R$ requested by Sm-GW $s = 1$ from operator $o = 1$ when inter-operator sharing is employed. Sm-GW $s = 1$ requests bitrate $R_{12} = 100 - R$ Mbps from operator $o = 2$. Also, Sm-GW $s = 2$ requests fixed bitrates $R_{21} = R_{22} = 20$ Mbps from each operator. Each operator o has a fixed uplink transmission bitrate constraint of $K_o = 50$ Mbps. Note that operator $o = (m) = 1$ has slack uplink transmission bitrate when $R \leq 30$ Mbps and can thus serve as roaming operator for the excess traffic to operator $e = 2$. As R increases and starts to exceed 70 Mbps, the roles are reversed, so that operator $e = 1$ can send some of its excess traffic to roaming operator $m = 2$.

Focusing initially on the case $R = 100$ Mbps, i.e., the right edge of Fig. 10, we observe that without SDN orchestrated optimization, Sm-GW $s = 1$ can only transmit at its fixed static allocation of $X_{11} = K_1/S = 25$ Mbps to operator $o = 1$, even though Sm-GW $s = 1$ has a traffic load demanding $R_{11} = R = 100$ Mbps. At the same time, Sm-GW $s = 2$ transmits at its requested rate $X_{21} = R_{21} = 20$ Mbps $< K_1/S$. Thus, the operator $o = 1$ uplink transmission bitrate K_1 is underutilized, even though Sm-GW $s = 1$ has more traffic to send, but cannot due to the inflexible static uplink transmission bitrate allocations.

With SDN orchestrated optimization with proportional sharing (but without inter-operator sharing), the overloaded uplink transmission bitrate $K_1 = 50$ Mbps of operator $o = 1$ is shared between the two Sm-GWs, allowing Sm-GW $s = 1$ to transmit $X_{11} = R_{11}/(R_{11} + R_{21}) = 41.7$ Mbps, while Sm-GW $s = 2$ transmits $X_{21} = R_{21}/(R_{11} + R_{21}) = 8.3$ Mbps. However, the uplink transmission bitrate K_2 of operator $o = 2$ is underutilized with only Sm-GW $s = 2$ transmitting $X_{22} = 20$ Mbps.

With inter-operator sharing, the unutilized uplink transmission bitrate $\zeta = K_2 - X_{22} = 30$ Mbps of operator $o = 2$, is used to carry excess traffic from operator $o = 1$. In particular, the aggregate of the regular operator $o = 1$ uplink transmission bitrate K_1 and the uplink transmission bitrate available to operator $o = 1$ through traffic sharing to operator $o = 2$ ($\zeta = 30$ Mbps), i.e., $K_1 + \zeta = 80$ Mbps is available to operator $o = 1$. With proportional sharing, Sm-GW $s = 1$ can transmit $X_{11} = (K_1 + \zeta)R_{11}/(R_{11} + R_{21}) = 66.7$ Mbps, while Sm-GW $s = 2$ can correspondingly transmit $X_{21} = 13.3$ Mbps, fully utilizing the backhaul capacities of both operators.

Overall, we observe from Fig. 10 that across the entire range of traffic loads R from Sm-GW $s = 1$ for operator $o = 1$, our SDN based multi-operator orchestration with sharing is able to fully utilize the uplink transmission bitrates of both operators. Note in particular, that depending on the traffic load, the roles of the two operators (excess or sharing) are dynamically adapted.

VI. Conclusion

We have introduced a new backhaul architecture by inserting a novel Smart Gateway (Sm-GW) between the wireless base stations (eNBs) and the conventional operator gateways, e.g., LTE S/P-GWs. The Sm-GW enables flexible support for large numbers of small cell base stations. In particular, the Sm-GW adaptively schedules uplink backhaul transmission grants to the individual eNBs on a fast (typically millisecond) timescale. In addition, an SDN orchestrator adapts the allocation of the uplink transmission bitrate of the conventional gateways of multiple operators to the Sm-GWs on a slow (typically minutes or hours) time scale. Simulation results have demonstrated that the scheduling of eNB grants by the Sm-GW can greatly improve the backhaul service over conventional static backhaul uplink transmission bitrate allocations. Moreover, the SDN orchestrator substantially improves the utilization of the backhaul bandwidth, especially when inter-operator sharing is permitted.

There are several important directions for future research on the Sm-GW architecture and small cell backhaul in general. One direction is to examine a variety of scheduling algorithms in the context of the Sm-GW. Another direction is to examine different specific optimization objective functions within the general SDN orchestrator optimization introduced in this article. Moreover, it is of interest to investigate QoS strategies for different traffic types, such as data, voice, and video traffic.

References

[1] A. H. Jafari *et al.*, "Small cell backhaul: Challenges and prospective solutions," *EURASIP J. Wireless Commun. Netw.*, vol. 2015, no. 1, pp. 1–18, 2015.

[2] C. Shen, J. Xu, and M. van der Schaar, "Silence is gold: Strategic small cell interference management using tokens," in *Proc. IEEE GLOBECOM*, Austin, TX, USA, Dec. 2014, pp. 4359–4365.

[3] W. Ni, R. P. Liu, I. B. Collings, and X. Wang, "Indoor cooperative small cells over Ethernet," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 100–107, Sep. 2013.

[4] U. Siddique, H. Tabassum, E. Hossain, and D. Kim, "Wireless backhauling of 5G small cells: Challenges and solution approaches," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[5] E. Mugume and D. K. C. So, "Spectral and energy efficiency analysis of dense small cell networks," in *Proc. IEEE VTC*, Glasgow, U.K., May 2015, pp. 1–5.

[6] F. Haider, C. Wang, B. Ai, H. Haas, and E. Hepsaydir, "Spectral/energy efficiency tradeoff of cellular systems with mobile femtocell deployment," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3389–3400, May 2016.

[7] S. Chen, J. Zhao, M. Ai, D. Liu, and Y. Peng, "Virtual RATs and a flexible and tailored radio access network evolving to 5G," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 52–58, Jun. 2015.

[8] T.-C. Liu, K. Wang, C.-Y. Ku, and Y.-H. Hsu, "QoS-aware resource management for multimedia traffic report systems over LTE-A," *Comput. Netw.*, vol. 94, pp. 375–389, Jan. 2016.

[9] K. Samdanis, R. Shrivastava, A. Prasad, D. Grace, and X. Costa-Perez, "TD-LTE virtual cells: An SDN architecture for user-centric multi-eNB elastic resource management," *Comput. Commun.*, vol. 83, pp. 1–15, Jun. 2016.

[10] M.-C. Chuang, M. C. Chen, and S. Yeali, "Resource management issues in 5G ultra dense smallcell networks," in *Proc. IEEE ICOIN*, Siem Reap, Cambodia, Jan. 2015, pp. 159–164.

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

[12] CISCO Systems Inc. (2014). *Cisco 4G LTE WWAN Deployment Guide*. [Online]. Available: http://www.cisco.com/c/dam/en/us/td/docs/routers/access/interfaces/software/deployment/guide/c07-731484-00-ngewan.pdf

[13] E. Menachi and R. Giladi, "Hierarchical Ethernet transport network architecture for backhaul cellular networks," *Wireless Netw.*, vol. 19, no. 8, pp. 1933–1943, 2013.

[14] Z. Ghebretensaé, J. Harmatos, and K. Gustafsson, "Mobile broadband backhaul network migration from TDM to carrier Ethernet," *IEEE Commun. Mag.*, vol. 48, no. 10, pp. 102–109, Oct. 2010.

[15] A. Mathew, T. Das, P. Gokhale, and A. Gumaste, "Multi-layer high-speed network design in mobile backhaul using robust optimization," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 4, pp. 352–367, Apr. 2015.

[16] H. Ali-Ahmad *et al.*, "An SDN-based network architecture for extremely dense wireless networks," in *Proc. IEEE Conf. SDN Future Netw. Services (SDN4FNS)*, Trento, Italy, Nov. 2013, pp. 1–7.

[17] S. Auroux, M. Dräxler, A. Morelli, and V. Mancuso, "Dynamic network reconfiguration in wireless DenseNets with the CROWD SDN architecture," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Paris, France, Jun. 2015, pp. 144–148.

[18] D. Wang, L. Zhang, Y. Qi, and A. U. Quddus, "Localized mobility management for SDN-integrated LTE backhaul networks," in *Proc. IEEE VTC*, Glasgow, U.K., May 2015, pp. 1–6.

[19] S. Zhang, C. Kai, and L. Song, "SDN based uniform network architecture for future wireless networks," in *Proc. Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Hefei, China, Jul. 2014, pp. 1–5.

[20] L. De la Oliva *et al.*, "Xhaul: Toward an integrated fronthaul/backhaul architecture in 5G networks," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 32–40, Oct. 2015.

[21] I. Elgendi, K. S. Munasinghe, and A. Jamalipour, "A three-tier SDN architecture for DenseNets," in *Proc. IEEE Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Cairns, QLD, Australia, 2015, pp. 1–7.

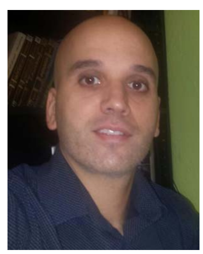
[22] J. Costa-Requena, V. F. Guasch, and J. L. Santos, "Software defined networks based 5G backhaul architecture," in *Proc. ACM Int. Conf. Ubiquitous Inf. Manag. Commun.*, Bali, Indonesia, 2015, Art. no. 35.

[23] S. Costanzo *et al.*, "An SDN-based virtual cell framework for enhancing the QoE in TD-LTE pico cells," in *Proc. IEEE QoMEX*, Pylos, Greece, 2015, pp. 1–6.

[24] M. Dräxler and H. Karl, "Dynamic backhaul network configuration in SDN-based cloud RANs," *arXiv preprint arXiv:1503.03309*, 2015.

[25] A. Hurtado-Borràs, J. Palà-Solé, D. Camps-Mur, and S. Sallent-Ribes, "SDN wireless backhauling for small cells," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., 2015, pp. 3897–3902.
[26] V. Jungnickel *et al.*, "Software-defined open architecture for front-and backhaul in 5G mobile networks," in *Proc. IEEE Int. Conf. Transp. Opt. Netw. (ICTON)*, Graz, Austria, Jul. 2014, pp. 1–4.
[27] C. Liu, K. Sundaresan, M. Jiang, S. Rangarajan, and G.-K. Chang, "The case for re-configurable backhaul in cloud-RAN based small cell networks," in *Proc. IEEE INFOCOM*, Turin, Italy, 2013, pp. 1124–1132.
[28] A. M. Akhtar, X. Wang, and L. Hanzo, "Synergistic spectrum sharing in 5G HetNets: A harmonized SDN-enabled approach," *IEEE Commun. Mag.*, vol. 54, no. 1, pp. 40–47, Jan. 2016.
[29] A. Gopalasingham *et al.*, "Generalized software defined network platform for radio access networks," in *Proc. IEEE CCNC*, Las Vegas, NV, USA, 2016, pp. 626–629.
[30] C. Niephaus, O. G. Aliu, M. Kretschmer, S. Hadzic, and G. Ghinea, "Wireless back-haul: A software defined network enabled wireless back-haul network architecture for future 5G networks," *IET Netw.*, vol. 4, no. 6, pp. 287–295, Nov. 2015.
[31] 3GPP. (2015). *Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification*. [Online]. Available: http://www.3gpp.org/dynareport/36321.htm
[32] S. Chaudhuri, I. Baig, and D. Das, "Utility based QoS aware uplink scheduler scheme for LTE small cell network," in *Proc. IEEE ICC*, London, U.K., Jun. 2015, pp. 3149–3154.
[33] S. Lakshminarayana, M. Assaad, and M. Debbah, "H-infinity control based scheduler for the deployment of small cell networks," *Perform. Eval.*, vol. 70, nos. 7–8, pp. 513–527, Jul. 2013.
[34] O. Semiari, W. Saad, S. Valentin, M. Bennis, and H. V. Poor, "Context-aware small cell networks: How social metrics improve wireless resource allocation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5927–5940, Nov. 2015.
[35] X. Wang, E. Visotsky, and A. Ghosh, "Dynamic cell muting for ultra dense indoor small cell deployment scenario," in *Proc. IEEE ICCW*, London, U.K., Jun. 2015, pp. 148–153.
[36] I. Gasparis, U. Kozat, and M. Sunay, "Programming flows in dense mobile environments: A multi-user diversity perspective," *arXiv preprint arXiv:1506.07816*, 2015.
[37] T. Taleb, Y. Hadjadj-Aoul, and K. Samdanis, "Efficient solutions for enhancing data traffic management in 3GPP networks," *IEEE Syst. J.*, vol. 9, no. 2, pp. 519–528, Jun. 2015.
[38] 3GPP. (2015). *Architecture Aspects of HomeNodeB(HNB)/Home Enhanced NodeB(HeNB)*. [Online]. Available: http://www.3gpp.org/DynaReport/23830.htm
[39] 3GPP. (2016). *TS 23.251 Universal Mobile Telecommunications System (UMTS), LTE, Network Sharing, Architecture and Functional Description, V13.2.0.* [Online]. Available: http://www.3gpp.org/DynaReport/23251.htm
[40] 3GPP. (2015). *General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access*. [Online]. Available: http://www.3gpp.org/DynaReport/23401.htm/
[41] A. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (SDONs): A comprehensive survey," *IEEE Commun. Surveys Tuts.*, to be published.
[42] Y. Yoshida *et al.*, "SDN-based network orchestration of variable-capacity optical packet switching network over programmable flexi-grid elastic optical path network," *J. Lightw. Technol.*, vol. 33, no. 3, pp. 609–617, Feb. 1, 2015.
[43] A. Blenk, A. Basta, M. Reisslein, and W. Kellerer, "Survey on network virtualization hypervisors for software defined networking," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 655–685, 1st. Quart., 2016.
[44] R. Mijumbi *et al.*, "Network function virtualization: State-of-the-art and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 236–262, 1st Quart., 2016.
[45] K. Nakauchi and Y. Shoji, "WiFi network virtualization to control the connectivity of a target service," *IEEE Trans. Netw. Service Manag.*, vol. 12, no. 2, pp. 308–319, Jun. 2015.
[46] R. Riggio, M. K. Marina, J. Schulz-Zander, S. Kuklinski, and T. Rasheed, "Programming abstractions for software-defined wireless networks," *IEEE Trans. Netw. Service Manag.*, vol. 12, no. 2, pp. 146–162, Jun. 2015.
[47] Wi-Fi Alliance. (2014). *Wi-Fi CERTIFIED Passpoint*. [Online]. Available: http://www.wi-fi.org/discover-wi-fi/wi-fi-certified-passpoint
[48] J. Zheng and H. T. Mouftah, "A survey of dynamic bandwidth allocation algorithms for Ethernet passive optical networks," *Opt. Switching Netw.*, vol. 6, no. 3, pp. 151–162, Jul. 2009.
[49] X. Bai, A. Shami, and C. Assi, "On the fairness of dynamic bandwidth allocation schemes in Ethernet passive optical networks," *Comput. Commun.*, vol. 29, no. 11, pp. 2123–2135, 2006.
[50] R. Jain, D. Chiu, and W. Hawe, *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System*, vol. 38. Eastern Res. Lab., Digit. Equipment Corp., Hudson, MA, USA, 1984.
[51] 3GPP. (2015). *TS 23.203: Policy and Charging Control Architecture*. [Online]. Available: http://www.3gpp.org/DynaReport/23203.htm

Akhilesh S. Thyagaturu received the B.E. degree from the Siddaganga Institute of Technology, Tumkur, India, in 2010, and the M.S. degree in electrical engineering from Arizona State University, Tempe, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering. He was with Qualcomm Technologies Inc., San Diego, CA, USA, as an Engineer from 2013 to 2015. He is also currently interning for Intel Mobile Communications Group, Intel Corporation, San Diego.



Yousef Dashti received the B.S., M.S.E., and Ph.D. degrees in electrical engineering with emphasis on signal processing and communications from Arizona State University, Tempe, USA, in 2003, 2008, and 2016, respectively. He is an Assistant Professor with the College of Technological Studies, Public Authority for Applied Education and Training, Kuwait. His research interests include fiber-wireless networking, optical networking, and long term evolution back-haul networks.

Martin Reisslein (A'96–S'97–M'98–SM'03–F'14) received the Ph.D. degree in systems engineering from the University of Pennsylvania in 1998. He is a Professor with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON EDUCATION as well as *Computer Networks* and *Optical Switching and Networking*.