OpenFlow-based Mechanisms for QoS in LTE Backhaul Networks

Luciano Jerez Chaves*, Islene Calciolari Garcia†, and Edmundo Roberto Mauro Madeira†

* Computer Science Department, Federal University of Juiz de Fora (UFJJ), Brazil.
† Institute of Computing, University of Campinas (Unicamp), Brazil.

Emails: {lchaves, islene, edmundo}@ic.unicamp.br

Abstract—The growing data traffic demand is forcing network operators to deploy more base stations, culminating in dense heterogeneous networks that require a high-connectivity backhaul. This scenario imposes significant challenges for current and future cellular networks, and Software Defined Networking (SDN) has been pointed as an enabling technology to overcome existing limitations. This paper shows how the OpenFlow protocol can be integrated into existing Long Term Evolution (LTE) networks to provide the required Quality of Service (QoS) in the network infrastructure. Three OpenFlow-based mechanisms are proposed: a traffic routing, an admission control function, and a traffic coexistence mechanism. Together, they can effectively control the bandwidth usage in the backhaul infrastructure, improving the QoS and ensuring a better user experience. Simulations were performed to validate the proposed mechanisms and highlight the benefits that can be achieved with the flexibility offered by the SDN technology.

I. INTRODUCTION

The world is witnessing a rapid growth in mobile communication. According to Cisco Systems, Inc [1], global mobile data traffic grew 69% in 2014, and it is expected to increase 10 times up to 2019. Not only the traffic is increasing, but also the number of connected devices. The 4th Generation (4G) mobile connections will grow at an annual rate of 46% for the next years. By 2019, 26% of all global devices and connections will be 4G-capable. Even though 4G connections only account for 2.9% of mobile connections, they already generate 30% of mobile traffic, mainly consumed by multimedia streaming. To accommodate all this traffic, Long Term Evolution (LTE) networks have been de facto employed for high-speed wireless communication, providing a highly efficient, packet-optimized service. However, the growing demand is forcing network operators to deploy more base stations to increase capacity through frequency reuse, especially in densely populated areas [2]. Also, the evolution toward 5G systems is converging to heterogeneous networks, bringing several wireless technologies together. A scenario like that requires a backhaul network with higher connectivity, thus, it is necessary to assist the access to these networks [3].

As indicated by the Open Networking Foundation (ONF), network operators are facing challenges as demand for mobility and bandwidth increases, and meeting current market requirements is virtually impossible with traditional network architectures [4]. Vendor dependence and lack of open interfaces limit the ability of network operators to tailor the network to their individual environments. Different solutions have emerged to this area, as the Network Function Virtualization (NFV) [5], which aims to use cloud technologies to consolidate network functions traditionally implemented in dedicated hardware. In addition, Software Defined Networking (SDN) and the OpenFlow protocol [4] have emerged as a promising paradigm, providing a shift toward a flow-centric model designed to enable more agile and flexible networks. In contrast to the current model, it decouples the control and data planes, allowing simplified forwarding hardware controlled by intelligent centralized software, which is assisted by applications that orchestrate service delivery in the network. By centralizing network intelligence, decision-making is facilitated based on a global (or domain) view of the network. Yazici et al. [6] advocate that a key differentiator of 5G systems will be in how to orchestrate the overall system control and SDN is a natural choice. This evolution is also debated by Sama et al. [7], where SDN is used as an outstanding enabler for NFV. By enabling these new technologies, it is viable to deploy more services with fewer costs, assuring them a long-term solution.

This paper discusses how the OpenFlow protocol [8] can be integrated to the existing 4G LTE networks. Three novel mechanisms that can be used to provide the required QoS in the LTE backhaul infrastructure are proposed. An OpenFlow-based traffic routing solution is used to exploit the SDN flow-based routing capability, allowing the use of different routing strategies by the centralized controller. A control admission mechanism is used to ensure that the available bandwidth is sufficient to accommodate new connection requests. Finally, a coexistence mechanism is proposed to limit traffic throughput. It ensures that best-effort traffic will not consume the bandwidth reserved for flows with strict QoS requirements. Furthermore, it exploits the OpenFlow queue support to improve Voice over IP (VoIP) experience in the network. Simulations are performed in the Network Simulator 3 (ns-3) to validate the proposed solutions, and the results confirm the effectiveness of these mechanisms.

The remaining of this document is organized as follows: Section II presents some background concepts on LTE networks, whereas Section III shows how the SDN paradigm can be integrated to the LTE technology, along with some related work on the topic. Section IV introduces the proposed QoS mechanisms and discusses their performance evaluation. The paper is concluded in Section V, summarizing results and outlining future works.
II. LONG TERM EVOLUTION (LTE)

LTE is a 3rd Generation Partnership Project (3GPP) collection of standards for high-speed wireless communication. It consists of the radio access and core networks, through the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) and Evolved Packet Core (EPC), respectively. Together, they comprise the Evolved Packet System (EPS), as illustrated in Fig. 1. The E-UTRAN is made up of essentially the Evolved Node B (eNB), which provides the air interface towards the User Equipment (UE). The EPC supports access to the Packet Data Network (PDN) domain only and consists of the PDN GateWay (P-GW) and the Serving GateWay (S-GW) for packet forwarding in the data plane. The P-GW provides connectivity from the UEs to the external packet data networks, handling IP addresses, packet filtering, and policy enforcement. The S-GW handles tunnels, mobility, and also acts as an anchor point for handovers. The eNBs are connected to the EPC by the S1 interface while the gateways are interconnected by the S5/S8 interface. Over these interfaces, the GPRS Tunnelling Protocol (GTP) over UDP/IP is used to encapsulate the traffic. In the control plane, the Mobility Management Entity (MME) is a key element, in charge of security, gateway selection, mobility, roaming, and handovers. The Home Subscriber Server (HSS) stores information for each UE, while the Policy Control and Charging Rules Function (PCRF) handle QoS policy decisions. The EPS uses the concept of bearers to route IP traffic over the EPC [9]. An EPS bearer uniquely identifies packet flows that receive a common QoS treatment. Each bearer is associated with a QoS Class Identifier (QCI) that specifies the resource type, scheduling priority, packet delay budget, and acceptable packet loss rate. Nine different QCI profiles are defined by 3GPP and are shown in Table I. The resource type can be classified into Guaranteed Bit Rate (GBR) bearers, which have an associated minimum bit rate value for which dedicated transmission resources are permanently allocated; and Non-Guaranteed Bit Rate (Non-GBR) bearers, which do not guarantee any particular bit rate. Services utilizing GBR bearers can assume that congestion-related packet losses (overflowing buffers) will not occur. When the UE attaches to the network a Non-GBR bearer is established, called the default bearer. Additional GBR or Non-GBR dedicated bearers can also be established at any time during or after completion of the attach procedure, accordingly to bearer-level QoS parameter values received by the P-GW from the PCRF. Since the EPS bearer has to cross multiple interfaces, the bearers are mapped onto lower layer bearers, as shown in Fig. 1.

III. SDN AND LTE INTEGRATION

A. Existing solutions in the literature

As discussed in Chaves et al. [10], the LTE architecture is complex and inflexible. The E-UTRAN provides high throughput under ideal conditions, but eNBs suffer from efficient coordination on dense deployments [11]. The EPC relies on specialized gateway equipment, which demand elevated reliability and, therefore, higher cost [12]. Several recent attempts have been done to optimize the existing mobile architecture.

In the mobile backhaul network, a congestion control solution is proposed by Venmani et al. [13], where the OpenFlow backhaul network can be shared among mobile operators, providing backup links in case of congestion. In the mobility management field, Gurusanthosh et al. [14] present the Semi-Distributed Mobility Anchoring (SDMA) mechanism, where the OpenFlow is used to distribute the mobility anchor points to reduces the signaling load in the backhaul network. Mahmoodi et al. [15] realize the SDN concept into the backhaul network through a distributed SDN control plane including the MME, HSS, and PCRF elements. They show that the new architecture reduces the UE power consumption and the signaling overhead between in backhaul entities.

### TABLE I

<table>
<thead>
<tr>
<th>QCI</th>
<th>Resource type</th>
<th>Priority</th>
<th>Packet delay budget (ms)</th>
<th>Packet error loss rate</th>
<th>Example services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GBR</td>
<td>2</td>
<td>100</td>
<td>$10^{-2}$</td>
<td>Conversational voice</td>
</tr>
<tr>
<td>2</td>
<td>GBR</td>
<td>4</td>
<td>150</td>
<td>$10^{-3}$</td>
<td>Conversational video (live streaming)</td>
</tr>
<tr>
<td>3</td>
<td>GBR</td>
<td>3</td>
<td>50</td>
<td>$10^{-3}$</td>
<td>Real-time gaming</td>
</tr>
<tr>
<td>4</td>
<td>GBR</td>
<td>5</td>
<td>300</td>
<td>$10^{-6}$</td>
<td>Non-conversational video (buffered streaming)</td>
</tr>
<tr>
<td>5</td>
<td>Non-GBR</td>
<td>1</td>
<td>100</td>
<td>$10^{-6}$</td>
<td>IMS signaling</td>
</tr>
<tr>
<td>6</td>
<td>Non-GBR</td>
<td>6</td>
<td>300</td>
<td>$10^{-6}$</td>
<td>Video (buffered streaming) TCP-based (HTTP, P2P, etc.)</td>
</tr>
<tr>
<td>7</td>
<td>Non-GBR</td>
<td>7</td>
<td>100</td>
<td>$10^{-3}$</td>
<td>Voice, Video (live streaming), Interactive gaming</td>
</tr>
<tr>
<td>8</td>
<td>Non-GBR</td>
<td>8</td>
<td>300</td>
<td>$10^{-6}$</td>
<td>Video (buffered streaming) TCP-based (HTTP, P2P, etc.)</td>
</tr>
<tr>
<td>9</td>
<td>Non-GBR</td>
<td>9</td>
<td>300</td>
<td>$10^{-6}$</td>
<td>Video (buffered streaming) TCP-based (HTTP, P2P, etc.)</td>
</tr>
</tbody>
</table>
In the mobile core network, the routing protocol is one of the most important issues, and many approaches keep using traditional tunnel-based traffic routing. This is the case of the MobileFlow, proposed by Pentikousis et al. [16], where the entire control plane is moved to software, and the datapath is composed only by simple elements with tunnel handling support. The Vertical Forwarding concept introduced by Hampel et al. [12] is also used to handle tunnels in mobile networks, using only a few switches that tunnel and de-tunnel the user-plane data. Their small number of switches keeps centralized control within scalability bounds, and remaining infrastructure can apply conventional routing protocols dictated by legacy domains. On the other hand, Jin et al. [17] introduce SoftCell, a disruptive solution that removes GTP tunnels and replaces all datapath EPC elements by simple OpenFlow switches and a set of middleboxes. The controller directs traffic over the network and middleboxes based on the traffic service. Kempf et al. [18] present a study on the evolution of cloud-based EPC, where all the control functions of EPC elements are moved into the cloud. The user plane is shifted into the OpenFlow switches, and these switches are extended to support GTP. In the traffic offloading context, Nagaraj et al. [19] introduce ProCel, where an OpenFlow switch at each eNB is used to classify traffic and act as the gateway for non-core traffic. It exploits the opportunity that most of the network traffic does not call for strict QoS requirements, and these flows are steered directly to fixed IP networking, considerably reducing the traffic on the EPC core, decreasing signaling overhead and improving scalability. A new OpenFlow-based control plane for EPC is presented by Said et al. [20], with the focus on resiliency and load balancing. The proposed architecture easily ensures the on-demand connectivity service even in critic situation such as network equipment failure and overload situations.

B. Proposed integration scenario

Instead of dismantling the LTE architecture, the proposed integration scenario focuses on the use of SDN at the backhaul network, connecting eNBs to the EPC via OpenFlow switches. The backhaul network is responsible for carrying all the traffic for S1 and X2 interfaces, which is encapsulated with the GTP/UDP/IP protocols. The proposed scenario preserves the GTP routing, so no EPS element that handles tunnel endpoints needs to be replaced and no changes in control protocols are necessary. This allows a soft integration of new OpenFlow network elements within a legacy infrastructure. The OpenFlow was extended with two new GTP match fields: the 2 byte GTP header flags field and the 4 byte GTP Tunnelling End ID (TEID) field. This approach dismisses the de/encapsulation operations, but can only match fields at tunnel headers (the outermost IP and UDP headers). In this way, the OpenFlow protocol is used to bring relevant information from encapsulated payload to tunnel headers.

Considering the network topology, Fig. 2 shows the wired portion for the proposed scenario. The backhaul network is a ring with an arbitrary number of OpenFlow switches. The gateway element is attached to the first switch, and eNBs are connected to the other switches. Connections are built with fast Ethernet full duplex links (100 Mbps), with an average delay estimated at 100 µs, for a typical 20 km metropolitan fiber cable. The ring was chosen as most of the legacy backhaul networks have a ring access topology, and it is one of the most efficient approaches in terms of protection and costs. In the access network, the eNBs are placed on a hexagonal grid with an inter-site distance of 500 meters. Each eNB covers an arbitrary number of active UEs, which are scattered close to it. This scenario was implemented in the Network Simulator 3 (ns-3), using the new OpenFlow 1.3 module (OFSwitch13), introduced by the same authors of this paper in Chaves et al. [10]. Table II summarizes the LTE parameters that were adjusted for the simulations.

As LTE networks are focused on VoIP and multimedia applications, four applications are used during the simulations to provide five different types of traffic flows in the network, with different QoS requirements: an HTTP application for web page access mapped to QCI 9; a VoIP conversation with call length expected to 100 seconds over UDP protocol and mapped to QCI 1; a live MPEG-4 video streaming over UDP protocol and mapped to QCIs 2 or 7; and a buffered MPEG-4 video streaming over TCP protocol and mapped to QCI 6 (videos with expected length to 90 seconds). UEs can request for randomly dedicated bearer context activations, with individual bearer requests following a Poisson Process with rate λ = 0.3 requests per minute.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>System frequency</td>
<td>2100 MHz</td>
</tr>
<tr>
<td>System bandwidth</td>
<td>20.0 MHz</td>
</tr>
<tr>
<td>eNB TX power</td>
<td>46 dBm</td>
</tr>
<tr>
<td>UE TX power</td>
<td>18 dBm</td>
</tr>
<tr>
<td>SRS periodicity</td>
<td>320 ms</td>
</tr>
<tr>
<td>Propagation model</td>
<td>OhBuildingsPropagationLossModel</td>
</tr>
<tr>
<td>MAC scheduler</td>
<td>CqafFMacScheduler</td>
</tr>
</tbody>
</table>
IV. QoS MECHANISMS FOR LTE BACKHAUL NETWORKS

The QoS mechanisms for LTE backhaul networks proposed in this paper are implemented in the OpenFlow EPC controller, which communicates with the MME element in order to monitor for EPS bearer context activation, modification, and deactivation procedures. The QoS mechanisms are related to network traffic routing, bearer admission control and GBR/Non-GBR traffic coexistence in backhaul infrastructure.

A. Traffic routing and bearer admission control mechanisms

1) Traffic routing: One of the advantages offered by the SDN paradigm is the possibility to exploit different routing strategies by the centralized controller. Once a new bearer is created in the network, the controller selects an acceptable route and install the match rules into the switches. Since backhaul traffic is encapsulated within GTP protocol, the header fields exposed to matching procedures are those related to the tunnel. To allow for a fine-grained per-flow routing, the OpenFlow switches consider only the GTP TEID identifier to route the traffic. In fact, this approach simplifies the match rules that are installed into switches, which speed up the lookup processes. For the ring network, the routing options are reduced to the clockwise or counter-clockwise direction. In this way, the controller can use two different routing policies:

- The Shortest Path Only, which always route the traffic over the shortest path between the source and destination;
- The Shortest Path First, which attempts to route the traffic over the shortest path whenever possible, or send the traffic in the other direction when necessary. This approach exploits the selective routing path advantage offered by the OpenFlow protocol.

2) Bearer admission control: The EPC controller also provides an admission control mechanism for GBR bearer activation and modification requests. To ensure the associated QCI requirements, GBR bearers call for some bandwidth allocation in the backhaul network. During the admission, the controller first checks for the requested bandwidth over all links in the selected routing path. When the requested bandwidth is available, the bearer activation/modification proceeds. Otherwise, the controller can check for a new routing path (if available), or block the bearer. When this happens, the MME is notified and the network aborts the context activation/modification procedure. The proposed mechanism monitors only for GBR bearer request. Once a bearer is blocked, the UE can optionally send the traffic over any Non-GBR bearer, with no QoS guarantees.

3) Performance evaluation: Simulations were performed to investigate how the traffic routing and the bearer admission control mechanisms can work together. The backhaul ring network was configured with 5 switches (switch 0 connected to the S-GW, and one eNB attached to each remaining switch in the clockwise direction). Thirty percent of the UEs are attached to eNBs 1 and 2, while seventy percent are attached to eNBs 3 and 4. This unbalanced configuration is intentionally chosen to underline the selective routing benefits that can be achieved with the OpenFlow. The admission control mechanism is configured to reserve up to 40% of the link bandwidth for GBR bearers.

Fig. 3a shows the average block ratio for an increasing number of UEs. The shortest path first routing policy significantly reduces the block ratio in this scenario, lowering the average block ratio by about 3.7 percentage points after 250 UEs, and increasing by 33% the number of accepted bearers. This performance is expected on this unbalanced load configuration, as the high number of UEs connected to eNBs 3 and 4 congests the links on this side of the ring. However, the controller can be aware of the congested links, and can selectively route exceeding traffic over the inverted routing path. This not only reduces the average block ratio but also balance the block ratio among all UEs in the network. This is illustrated in Fig. 3b, which shows the blocked requests and the average routing distribution among the shortest and inverted path for both routing policies, considering all 4 eNBs in the simulations for 250 UEs in the network.

B. GBR/Non-GBR coexistence mechanism

1) Motivation: The available resources in the LTE backhaul infrastructure network are shared among all GBR and Non-GBR bearers and the network is supposed to allocate the requested bandwidth for GBR bearers accepted by the admission control mechanism. As the traffic management is not specified by 3GPP standards, vendors and network operators can establish policies and set up equipment accordingly.

Some simulations were used to investigate how the GBR traffic can be affected by Non-GBR traffic in the backhaul network with shared bandwidth resources and no traffic management. The backhaul ring network was configured with 10
OpenFlow switches, with UEs equally distributed among 9 eNBs. The routing policy in use is the *shortest path first*, which exploits the selective routing. The admission control mechanism checks for the available bandwidth and blocks requests that exceed 40% of the total link bandwidth. Despite the admission control, there is no other mechanism in the backhaul network to assure resource reservation, and all packets traversing the network are handled in the same way, over the same single output queues.

Fig. 4 shows the GBR aggregated traffic throughput and packet loss ratio for an increasing number of UEs. As network traffic load grows, it is possible to observe an increasing gap between the GBR bandwidth reserved by the admission control and the actual GBR traffic throughput (Fig. 4a), mainly caused by the packet loss ratio (Fig. 4b). This increasing loss ratio is caused by overflowing buffers. In fact, when there is no available bandwidth to carry all network traffic, only Non-GBR packets should be discarded, as services over GBR bearers can assume that congestion-related packet losses will not occur. To overcome this situation, the coexistence mechanism has to keep both traffics within its allowed bandwidth.

2) Coexistence mechanism: The backhaul infrastructure can be built using L2 or L3 technologies. Many mobile operators want to keep their networks simpler by avoiding L3 protocols. Without routers, the LTE QoS mechanisms can be implemented at MAC layer as VLAN with traffic class priorities. Another commonly used approach is the MultiProtocol Label Switching (MPLS). However, as LTE is IP-based, it can leverage L3 Differentiated Services (DiffServ) [21] for providing QoS on a per-hop basis. DiffServ uses the 6-bit DiffServ Code Point (DSCP) field in the IP header for traffic forwarding treatment of each packet, based on the DSCP value [9]. To determine the per-hop traffic forwarding treatment of each individual packet, based on the DSCP value [9].

To enable for traffic separation and fast classification, OpenFlow switches connected to the gateway or to the eNBs implement a QCI to DSCP mapping function, translating from bearer-level QoS (QCI) to transport-level QoS (DSCP). In this way, packets associated with a specific QCI are marked with a specific DSCP for forwarding in the backhaul network [9]. The LTE QCI to DSCP mapping can be performed based on network operator policies, and is implemented here as indicated in Table III. OpenFlow switches can implement queue management schemes and scheduling algorithms to determine the per-hop traffic forwarding treatment of each individual packet, based on the DSCP value [9].

To limit traffic throughput up to the allowed bandwidth, the OpenFlow meter table is used to measure and control the rate of packets. The meter triggers a meter band if the packet rate or byte rate passing through the meter exceed a predefined threshold. Two types of meters are in use:

- **GBR per-flow meter**, used to limit the GBR bearer throughput within its maximum bit rate. These meters are installed together with the DSCP marking process.
- **Non-GBR coexistence meter**, used to limit the aggregated Non-GBR traffic throughput over each link, effectively reserving the bandwidth for active GBR bearers. These bearers are regularly updated by the admission control mechanism for newly accepted bearers.

3) Performance evaluation: Fig. 5 shows simulation results for the coexistence mechanism. It is possible to observe that Non-GBR coexistence meters can properly limit the Non-GBR traffic, allowing GBR traffic to use the reserved bandwidth. Fig. 5a shows the traffic throughput for both resource types while Fig. 5b highlights the GBR throughput with and without the coexistence mechanism (in comparison to Fig. 4a). Also, the packet loss ratio at queues (Fig. 5c) is reduced for both traffics, but a new packet drop ratio is introduced by the use of OpenFlow meters (Fig. 5d). It is possible to observe the constant GBR drop ratio triggered by the per-flow meters limiting the individual bearer throughput, and an increasing drop ratio due to Non-GBR coexistence meters.

4) VoIP improvement: An additional improvement to the coexistence mechanism is the use of OpenFlow queue support to better treat VoIP traffic. To provide the required voice quality, QoS capability must be added to the traditional data-only network. In an environment of mixed real-time and bulk...
traffic, it is natural to use priority queuing to provide the real-time traffic priority service. New simulations were performed, now mapping VoIP traffic to a second high-priority queue at each OpenFlow output port. Fig. 6 compares both delay and jitter for VoIP traffic over the backhaul network, considering the first simulations without traffic management (Coex. OFF), the simulations with the coexistence mechanism (Coex. ON), and this new group of simulations including the VoIP queuing mechanism (VoIP queues). The use of high-priority queues for VoIP traffic results in a constant average delay and jitter for this type of traffic, regardless how congested the network is. Also, it is possible to observe that, regardless priority queues, the coexistence mechanism also improved network average delay, with a small jitter increase for high load scenarios.

V. CONCLUSIONS AND FUTURE WORK

This paper explores how the OpenFlow protocol can be used to realize the LTE QoS requirements into mobile backhaul networks. Three OpenFlow-based mechanisms were introduced: traffic routing, bearer admission control, and the traffic coexistence. They were evaluated by means of simulations, and the results confirm the effectiveness of the solutions. For the routing and bearer admission control, the number of accepted bearers was increased by 33%, while the coexistence mechanism properly limited the Non-GBR traffic and allowed a constant delay and jitter for VoIP traffic, regardless how congested the network is. As future work, the authors aim to improve the mechanisms to support UE mobility.

REFERENCES


