Enabling Flexibility of Traffic Split Function in LTE-WiFi Aggregation Networks Through SDN

Suzan Bayhan and Anatolij Zubow
Telecommunication Networks Group (TKN), Technische Universität Berlin, Germany
{bayhan, zubow}@tkn.tu-berlin.de

***Abstract*—Unlicensed spectrum is very appealing for LTE operators to expand their capacity cost-effectively. Hence, they have recently been exploring various capacity-expansion approaches using unlicensed spectrum. One option, known as LTE-WiFi aggregation (LWA), lets an LTE eNB to deliver some of its traffic through the carrier WiFi APs so that an LTE user can get its downlink traffic from both interfaces. However, LWA requires careful splitting of the traffic between LTE and WiFi resources to avoid one link becoming congested or to ensure good load balance among the colocated network nodes, e.g., eNBs and WiFi APs. Currently, the eNB is in charge of deciding how to split the traffic. However, due to its local knowledge, the split may not be optimal. Instead, we take benefits of Software-Defined Networking (SDN) by decoupling the control and data forwarding plane, where a centralized entity, *traffic split function controller* (TSFC), using its broader knowledge can allocate resources more efficiently. We envision that a dumb datapath element (the TSF) splits flows/packets between ports, i.e., eNB and AP(s), as defined by a remote control process, i.e. the TSFC. However, it is unclear where this centralized entity should be located in the network hierarchy, e.g., at each eNB or deep in the core network or even in the cloud, due to several opposing forces. Centralization at the higher layers in the network hierarchy brings more global knowledge at the expense of increased delay and control message overhead. In this paper, we first list the required new interfaces between network elements for the envisioned new LWA architecture and next present a simple model to decide on the TSFC placement capturing various dynamics of the user, its traffic, and radio links.**

I. INTRODUCTION

A key challenge of the cellular operators nowadays is to meet the unprecedented increase in demand for mobile data traffic while keeping the costs low [1]. Desire for low-cost solutions has resulted in operators to focus on the unlicensed spectrum and already-available ubiquitous WiFi infrastructure. The common approach is to aggregate LTE with WiFi resources at different layers: unlicensed LTE at the modem level [2], LTE-WiFi aggregation (LWA) [3] at the radio access level, or classical offloading approach where aggregation takes place at the transport layer using Multipath TCP [4]. Unlicensed LTE uses the license-exempt spectrum that WiFi operates at, whereas LWA uses the WiFi infrastructure simultaneously with the LTE network. Hence, while the former has to overcome the coexistence challenge with the incumbent WiFi networks operating at 5 GHz UNII band, LWA can be realized without much hassles, however, with possibly lower performance gains compared to the former [5]. LWA needs only software updates for end user devices and network elements, e.g., eNB and WiFi APs, to support LWA. Similar to LTE dual-connectivity [6] defined in 3GPP specification Release 12, LWA facilitates a UE to be anchored to multiple network nodes, i.e., eNB and WiFi AP.

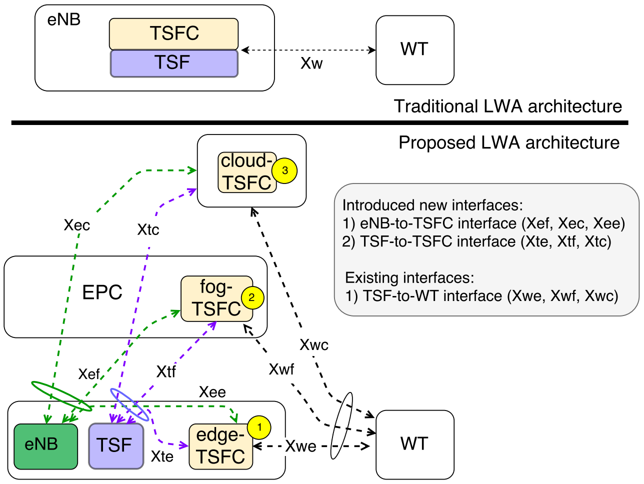


Fig. 1. Conventional LWA architecture (top figure) and proposed architecture (bottom figure). Possible locations to deploy traffic split function controller (TSFC), 1: edge-TSFC colocated with the eNB, 2: fog-TSFC, and 3: cloud-TFSC. Note that each interface (i.e., Xe, Xt, Xw) has three instances in the proposed architecture, however their functionality is the same.

In LWA, an eNB implements the *traffic split function* (TSF) and decides on whether to split and how to split the traffic. However, we argue that by decoupling the control function, *traffic split function controller* (TSFC), from the actual splitting function TSF, we can introduce flexibility as well as higher efficiency due to exploiting centralization gains. As discussed in [7], Software-Defined Networking (SDN) can introduce flexibility to the network, simplicity to the data-plane network elements, and performance gains to the provided services by decoupling control plane from the data plane performing actual forwarding. Besides, our motivation for moving traffic split functionality beyond the eNB is due to the fact that a WiFi AP might be connected to several other eNBs and its resources might be used by these eNBs. In that case, although an eNB expects that its user will get some promised rate from the assigned AP, this expected link capacity may not be achieved if other eNBs also decide to offload some of their traffic to the same WiFi AP. Hence, local decisions made by each eNB might affect performance of others. To cope with this discrepancy, the TSFC can be moved to deeps in the network, e.g., to the fog or to the cloud. However, such a change in the architecture introduces various challenges, e.g., additional control plane overhead as well as the latency between the actual split function and the split control function, i.e., TSFC. In this paper, we discuss first how the split function might be affected by this uncoupling, then list the required new interfaces between network elements, and finally present a simple model to decide on the TSFC location capturing various dynamics of the user, its traffic, and radio characteristics. While SDN-assistance for LWA is first introduced in [8], our work differs from [8] in that our goal is not to design the traffic split algorithm, but rather to explore the impact of the location of the controller.

Our contributions are two-fold:

- We introduce a novel LWA architecture as in Fig. 1 which benefits from SDN approach by decoupling the control and the actual traffic splitting in LWA to achieve more flexible and efficient (e.g., higher centralization gain) solutions. As this decoupling requires communication among the decoupled components, we introduce the necessary interfaces needed for the proposed flexible architecture.
- We identify the key parameters affecting the decision on where to deploy the TSFC; at the edge, the fog, or the cloud.

The rest of the paper is organized as follows. Sec. II overviews the key points of LWA while Sec. III describes a simple LWA setting. Next, Sec. IV introduces the proposed architecture. The following sections, Sec. V and Sec. VI, then detail how TSFC works depending on its location and provide a heuristic to decide on where to deploy TSFC. To develop an understanding of the proposed heuristic, Sec. VII investigates where the TSFC will be deployed under various mobility and flow scenarios. Sec.VIII summarizes the related work. Finally, Sec.IX concludes the paper with some future directions.

II. BACKGROUND ON LWA

LWA [1], [9] is one of the proposals toward expanding LTE networks' capacity exploiting the WiFi networks. In LWA, the key idea is to use the already-deployed carrier WiFi infrastructure (or third party WiFi networks) to carry some of the traffic originally to be served by the eNBs. In other words, LWA combines the radio resources of both LTE and WiFi at the packet data convergence protocol (PDCP) layer without requiring major changes to the WiFi and the LTE network. That is, LTE eNB can route the PDCP packet data units (PDUs) of the same IP flow independently using LTE and WiFi links [5]. In LWA, LTE eNB splits the traffic and sends some of the user traffic through the WiFi APs. We call this function *traffic split function* (TSF) and the logic to control how TSF works *traffic split function controller* (TSFC). A dual-mode user equipment, i.e., equipped with both LTE and WiFi interfaces, can simultaneously receive traffic on both interfaces. As opposed to LTE in unlicensed bands (i.e., LTE-U or LAA), LWA does not exhibit a challenge for WiFi in terms of coexistence. In 3GPP Release 13, LWA uses WiFi links only for the downlink (DL) traffic.

LWA supports two deployment scenarios: colocated and non-colocated. In the former, eNB and WiFi AP are encapsulated in the same unit or connected over an ideal link with negligible delay, whereas in the latter they are connected over a non-ideal backhaul using an interface called Xw interface. Xw supports both the data and control plane. UEs should support WiFi measurements, e.g., RSSI, channel utilization, or station count [1], [10], so that they can feed this information to the eNBs for a WiFi-aware traffic splitting.

To enable Xw communication without changing every single WiFi AP, LWA introduces new logical nodes called Wireless/WLAN Termination (WT) nodes which can provide the communication between the eNB and the WiFi AP. A WT can be connected to multiple eNBs and also to multiple WiFi APs [10]. It reports statistics for each of the connected WiFi APs to the eNB. The report includes AP load and available channel utilization. UEs also need to become LWA-aware to handle out-of-order packets due to the fact that both interfaces may experience different packet latencies. That is, a UE has to implement PDCP reordering and PDCP aggregation. This functionality can be added via a software update [9]. 3GPP Release 13 supports UEs to send feedback directly to eNB on the WLAN link performance via LWA status report to avoid dependency on AP status reports.

III. SYSTEM MODEL

We consider a setting consisting of an LTE network colocated with carrier WiFi APs. Each eNB is connected via Xw interface to a WT which controls one or more APs. Note that WT is a logical node and can be located in an access controller (AC), an AP, or as a stand-alone node in the WiFi network [9].

We assume that there is a fog compute environment in the proximity of a group of eNBs and this environment can host some of the control plane logic. The network link between the fog and corresponding TSF at eNBs has a delay of τ_f msecs. Moreover, the network operator can deploy some network functions at a cloud provider's site. We denote the delay between a TSF at eNB and the cloud by τ_c msecs. We assume the existence of a TSFC which can be deployed at one of the three locations in the network: at each eNB co-located with TSF, in each fog environment, or in the cloud environment. We will refer to each case by *edge-TSFC*, *fog-TSFC*, and *cloud-TSFC*, respectively. Note that actual data plane splitting is still performed at the eNB. That is, TSF is at the eNB, however the decision can be made by the controller, i.e., TSFC, which might be colocated with TSF or could be elsewhere in the network, i.e., fog or cloud. For an edge-TSFC, the TSF and TSFC is colocated.

There are multiple user equipments (UE) to be served by the network. Moreover, a UE moves with a constant speed of v kmph. As LWA supports currently DL data delivery, we consider only the DL traffic. Each UE has one or more traffic flows. The flows are assigned to bearers depending on their quality-of-service requirements, e.g., similar flows are considered as a single bearer. A TSFC makes the mode assignment at the granularity of a bearer. That is, it assigns each bearer a mode from the following two modes: *LTE-only* and *LWA* modes. In case of the former, packets belonging to the bearer can only follow the LTE link, whereas in the latter, both LTE and WiFi links are used. How to split the bearer traffic into two links depends on the *traffic split scheduling* policy. This decision can be made at packet or flow granularity.

In our model, each TSF stores a Flow Match Table (FMT) to look-up the rules for traffic splitting and decide on how to route a flow. FMT stores rules, e.g., bearer ID, destination UE ID, bearer mode. Note that FMT can be implemented using OpenFlow tables [11]. Similarly, a TSFC stores an FMT for the traffic controlled by it. The entries are: eNB ID, bearer ID, destination UE ID, bearer mode.

IV. PROPOSED SDN-BASED FLEXIBLE LWA ARCHITECTURE

It is almost a consensus in the networking community that networks need more flexibility at the core and less complexity at the edge due to the growing pressure on network operators to serve bandwidth-demanding applications/users cost-effectively. To this end, SDN proposes to move the intelligence towards the core which lets edge devices (e.g., eNBs) to be simpler and cheaper. In spirit of this approach, we propose to decouple TSFC from TSF for an LWA network. More specifically, we introduce a flexible LWA architecture (bottom figure in Fig. 1) which lets TSFC to be deployed at different locations in the network.

The flexibility of our proposal comes at the expense of two new interfaces that are needed for communication between eNB and TSFC as well as between TSF and TSFC. As shown in Fig. 1, we introduce the following two interfaces:

- **Xe interface** handles the control plane communication between eNB and the corresponding TSFC. For example, the statistics about eNB links to the users, traffic load are communicated to the TSFC through this channel. We refer to three instances of this interface as Xee, Xef, Xec, which correspond to the interface in the existence of edge-TSFC, fog-TSFC, and cloud-TSFC respectively.
- **Xt interface** handles the control plane communication between TSF and TSFC, e.g., SDN OF/P4, which we detail in the next section.

Note that we denote the Xw interface with Xwe, Xwf, Xwc to distinguish the three different controller placement options. Depending on the TSFC location, TSF and TSFC communications entail different delays on each interface. We denote the delay on each interface by $\tau_\star$ msecs where $\star$ is the name of the interface, e.g., ee, wc.

V. TRAFFIC SPLIT FUNCTION CONTROLLER (TSFC)

This section first introduces how TSF works and TSFC functions, then details the operation of TSF depending on TSFC location. Please recall that in our model, TSF can be a simple (software) switch and is co-located with the eNB whereas TSFC, whose location is flexible, has high processing power which enables it to implement sophisticated logic.

A. TSF

Fig. 2 illustrates the steps taken by the TSF upon a flow arrival. First, it checks the FMT to find a matching rule for this flow. If the flow matches a rule, then TSF takes the corresponding action defined by the rule, e.g., deliver the flow through LTE link. Otherwise, it applies the default rule. Note that FMT is populated via interaction with the TSFC. In the next section, we explain when this interaction takes place.

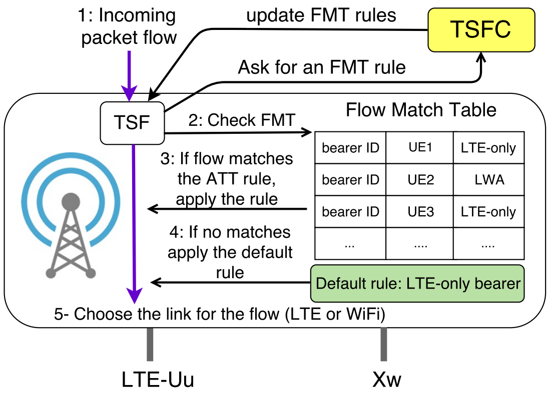


Fig. 2. Traffic split function. Upon arrival of a flow, TSF looks up FMT to check if there is a rule on which link to route this flow.

B. TSFC

TSFC has two functionality: bearer mode assignment and split scheduling for LWA bearers. Before explaining these functionality, let us first introduce two types of TSFCs: *reactive* or *proactive*.

- **Reactive TSF (rTSFC)**: In this case, the eNB triggers the rTSFC whenever a new flow arrives to the eNB similar to the *Packet-in* message in OpenFlow [11]. Based on the TSFC decision, eNB applies the policy. In rTSFC, the FMT update step takes place between Step 1 and Step 2 in Fig.2.
- **Proactive TSF (pTSFC)**: In this case, eNB interacts with the TSFC only periodically. In pTSFC, the FMT update step in Fig.2 takes place at the beginning of the TSFC period. Hence, the eNB must define a default rule on how to handle the traffic for which there is no matching rule in the FMT. TSFC period is a design parameter which can be tuned depending on the network dynamics, e.g., traffic inter-arrival time, and desired degree of prompt action.

Mode assignment scheme decides on whether to use only the LTE link or both LTE and WiFi radio links. We refer to the former as *LTE-only mode* and the latter as *LWA mode*. In case of rTSFC, when a new bearer is created, TSFC is triggered to run its mode assignment function. However, mode assignment can be performed periodically to adapt to the dynamics of the changes in the network, e.g., change in user locations, as well as the traffic flows, e.g., completed flows. In case of pTSFC, at the beginning of TSFC period, the pTSFC assigns a mode for each active bearer. The eNB applies its default rule for bearers that are created between two TSFC periods. Note that each TSFC selects an action, i.e., either mode assignment or traffic split scheduling, depending on its knowledge on the system resources. For flows whose duration are short compared to the change in the coherence time of the UE's LTE and WiFi links, rTSFC can exploit the best link compared to a pTSFC which is triggered earliest at the beginning of the next TSFC period.

After each bearer gets its bearer mode, the eNB knows how to deliver traffic of a flow belonging to the LTE-only bearers. On the other hand, for a flow in an LWA bearer, eNB needs information on how to split the traffic. Traffic split scheduling can be performed in the following two ways.

- **Per-flow traffic split**: TSFC makes the decision in the granularity of a flow. In other words, it assigns either WiFi or LTE for each flow within a bearer. As a result, unexpected transport layer issues (e.g., when using TCP) because of out-of-order packets are avoided as all packets belonging to a single flow follow only one particular link, either LTE or WiFi.
- **Per-packet traffic split**: TSFC makes the decision in the granularity of a packet, e.g., decides on the weight of each radio link to route a packet. In this case, packets belonging to the same flow might follow different links. As a result, users might enjoy throughput improvement even for a single flow. On the other hand, above-mentioned TCP-related problems may occur.

Note that the above approaches result in different storage requirements at TSF and TSFC as a result of different FMT sizes. The storage requirement will be lower for flow-level splitting. However, it may not achieve as high capacity as the packet-level traffic splitting.

C. Mode Assignment depending on the TSFC location

For rTSFC, eNB triggers the TSFC every time a new flow is created. Such an interruption might be exhaustive for a cloud-TSFC which in some cases has a control over the whole network to achieve the highest centralization gains. Moreover, latency of control plane communication between the TSF located in the eNB and cloud-TSFC might be high making it impractical to react to short lived flows or high channel dynamics due to mobility. Another reason for avoiding rTSFC is high mobility, e.g., vehicular scenarios. Hence, pTSFC is a better option for cloud-TSFC. For fog-TSFC, the degree of the challenge depends on the number of eNBs each fog controller is connected to. Note that determining the number of controllers is another decision the network provider needs to make considering various factors, e.g., network topology or performance goal [7].

In case of pTSFC, the TSF colocated with eNB receives some rules from the pTSFC only periodically. Hence, it may not have a rule already set by the pTSFC for some of the bearers. For such bearers, eNB has to apply its default rule, e.g., the LTE-only mode. While there is no straightforward way to decide on the pTSFC period or the TSFC location, one must consider the trade-off between scalability and the centralization gain one would achieve by moving the TSFC up in the network hierarchy.

D. Split scheduling depending on the TSFC location

For LWA bearers, TSF in eNB has to apply the splitting based on the TSFC decision on how to split the traffic. In rTSFC, TSFC considers the capacity of both WiFi and LTE links to decide on the best link to route the current flow. In case of pTSFC, this decision is given by the TSF in eNB: it applies its default rule for LWA traffic splitting if no specific rules exist in the FMT, otherwise applies the rule defined in the FMT. In the next TSFC period, all active flows are considered while making a decision on bearers and the traffic splitting. Similar to the mode assignment, using rTSFC might be impractical for cloud-TSFC or fog-TSFC. Hence, we believe that pTSFC should fit better to cloud-TSFC.

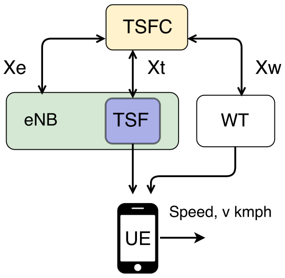


Fig. 3. Interfaces to TSFC from eNB, TSF, and WT. Recall that delay of each interface depends on the location of TSFC (i.e., edge, fog, cloud).UE speed is v kmph.

VI. WHERE TO DEPLOY TRAFFIC SPLIT FUNCTION CONTROLLER (TSFC)?

The optimal decision for mode assignment and traffic split scheduling decision depends on the location of the controller. Therefore, we focus on the controller placement problem in this paper (refer to Fig. 3). The controller should be able to react to changes in the network promptly. For this reason, we need to identify the key parameters resulting in a change in the user's radio link conditions. Below, we introduce several approaches which aim to react to changes in different scales, e.g., small scale changes as in the change of radio channel characteristics or longer scale changes such as user handovers.

A. TSCF placement considering small scale changes

Two insights that drive our heuristic are as follows:

- **Channel coherence time** τ_{ch}: For the TSFC decision be relevant for a UE, the time for communication between all interfaces in two ways, i.e., round-trip-time (RTT), must be shorter than the radio channel coherence time. Otherwise, the information that a remote-TSFC (i.e., fog-TSFC and cloud-TSFC) uses for its bearer mode and split scheduling decision becomes outdated. As a result, TSFC decision may result in performance loss compared to that which could be achieved by an edge-TSFC with a more up-to-date knowledge of the channel state.
- **Flow duration** (T_f): Similarly, duration of a flow is important as the decision about the flow has to be completed before it is finished.

With the above-listed insights, we define the *controller delay budget* ($\tau_{\max}$) as follows:

$$\tau_{\max} = \min(\tau^{l}_{ch}, \tau^{w}_{ch}, T_f) \text{ msecs}, \tag{1}$$

where τ^l_{ch} and τ^w_{ch} represent channel coherence time for LTE link and for WiFi link, respectively. We can approximate the coherence time of a radio channel with center frequency f and for a mobile user with speed v as:

$$\tau_{ch} \approx \frac{c}{fv} \text{ msecs,} \tag{2}$$

where c is the speed of light.

After calculating controller delay budget, we can decide on the location of the controller simply by comparing the RTT of each controller placement option. RTT delay of communication in a particular placement decision equals:

$$\tau_\star = 2\max(Xw\star, Xe\star, Xt\star), \tag{3}$$

where $\star \in \{e, f, c\}$ corresponding to edge, fog, and cloud, respectively.

Since operators are interested in minimizing their operational costs, our goal is to realize the highest centralization gain which translates into lowest cost for the operator. Hence, given this objective, we favor cloud-TSFC over fog-TSFC, and fog-TSFC over edge-TSFC. Hence, we propose the following heuristic to decide on the TSFC location:

$$TSFC = \begin{cases} cloud - TSFC & \text{, if } \tau_{\max} \geqslant \tau_c \\ fog - TSFC & \text{, if } \tau_f \leqslant \tau_{\max} \leqslant \tau_c \\ edge - TSFC & \text{, ow.} \end{cases} \tag{4}$$

B. TSCF placement considering shadowing

Rather than considering small scale changes due to multipath propagation in the user's radio channel, we can alternatively consider medium and longer scale changes. For the former, we consider the channel decorrelation time due to shadowing whereas, for the latter, we consider the user's handover from one cell to another.

Decorrelation time is simply the decorrelation distance d_{sh} divided by the user's speed: $\tau_{sh} = d_{sh}/v$ seconds. The decorrelation distance measures the distance where the signal autocorrelation equals $1/e$ of its maximum value [12]. This distance is on the order of blocking object's size and hence depends on the environment. Generally speaking, it is shorter for urban environment compared to sub-urban or rural environment. Then, we calculate the controller's delay budget as: $\tau_{\max} = \min(\tau_{sh}, T_f)$. We decide on the best TSFC placement based on (4).

C. TSCF placement considering the handovers

In Release 13, an LWA configuration is updated when the user handovers or from one eNB to another. Moreover, the user might handover from one WiFi AP to another upon change of its location, which might require an LWA reconfiguration. Hence, the time to handover becomes a constraint in the TSFC location. We can approximate the time to handover for a given user speed, the user's location in a cell, and the cell radius, as in [13]. For the sake of completeness, let us summarize the mentioned approach. Assume that the coverage radius of cell is r. The user's distance from its serving base station (i.e., WiFi AP or LTE eNB) is $a \sim U(0, r)$ and the user moves with a speed $v \sim U(v_{\min}, v_{\max})$ toward the cell edge with an angle $\theta \sim U(0, 2\pi)$. We find the expected time for a user who is a meters away from its base station to reach the cell edge—where the user-to-AP link quality is poor and hence a handover needs to be performed. To calculate this handover time T_{ho}, we first derive the expected distance from the cell exit point d_{ho} using the cosine theorem: $a^2 + d_{ho}^2 - 2 \cdot a \cdot d_{ho} \cdot \cos(\pi - \theta) = r^2$. We reorganize the above equation as a quadratic univariate equation where the only unknown is d_{ho}. Then, we can find the root of the above equation as: $d_{ho} = -a \cdot \cos\theta \pm \sqrt{(a \cdot \cos\theta)^2 + (r^2 - a^2)}$. Given this distance d_{ho}, we can derive time to handover $T_{ho} = d_{ho}/v$ which denotes the expected time to span d_{ho} with speed v. We calculate the time to handover for the LTE connection (τ^l_{ho}) and for the WiFi connection (τ^w_{ho}) separately. Then, we calculate the controller's delay budget as: $\tau_{\max} = \min(\tau^w_{ho}, \tau^l_{ho}, T_f)$ msecs. Finally, we apply our heuristic in (4) to decide on TSFC placement.

TABLE I
SIMULATION PARAMETERS.

Parameter	Value
Cell radius	LTE=166 m, WiFi=20 m
Frequency	LTE=2.3 GHz, WiFi={2.4/5 GHz}
Flow duration	Phase type distribution [14]
User speed	Stationary, low, medium, high [15]
Shadowing decorrelation distance model	U(0,4.38) meters [16]
X⋆e interface delays	5 ms
X⋆f interface delays	10 ms
X⋆c interface delays	50 ms

VII. PERFORMANCE EVALUATION

In this section, we analyze how our heuristic locates the TSFC under various settings via Monte-Carlo simulations on our custom Python simulator. To model the flow duration, we use the CDF function presented in [14] which analyzes a large scale dataset collected from a mobile network operator in China to identify the distribution of flow duration. Authors [14] fit their dataset with a Phase Type CDF and find the distribution parameters. Since the proposed distribution fits the collected data with high accuracy, we generate our flow duration using the suggested CDF and parameters in [14].

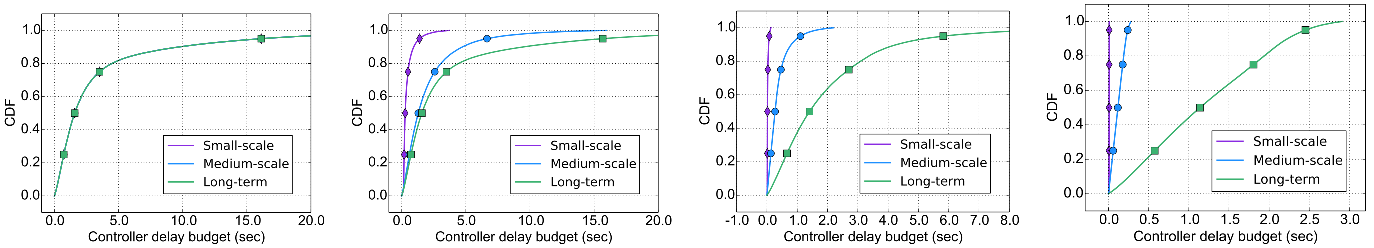
We consider the following mobility cases defined in [15]:

- no: static users (v=0),
- low: pedestrians where $v \in (0, 3]$ kmph,
- medium: slow moving vehicles where $v \in (3, 50]$ kmph, and
- high: fast moving vehicles where $v > 50$ kmph. For this scenario, we consider user speeds in $[50, 80]$ kmph.

We set LTE frequency to 2.3 GHz. We use the decorrelation distances suggested by [16]: for an urban environment, for a LOS link, the decorrelation distance is 4.25 meters and 4.5 meters for OLOS. Then, to model the distribution of the decorrelation distance, we consider the average of LOS and OLOS distances and use the following uniform distribution variable $U(0, 4.38)$ meters.

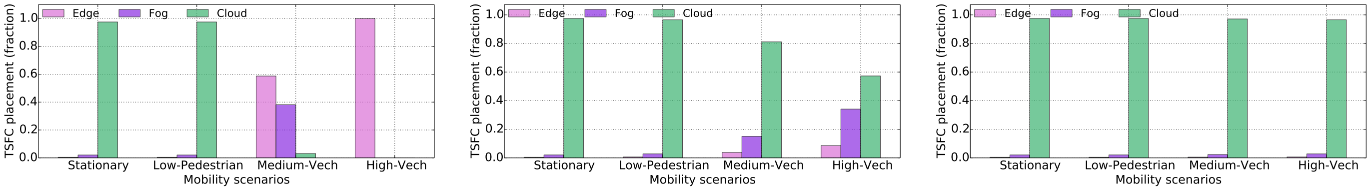
A. Edge, Fog, or Cloud: where to deploy the TSFC?

Using our approach in Sec.VI, we present some recommendations on the TSFC location for different scenarios. We find

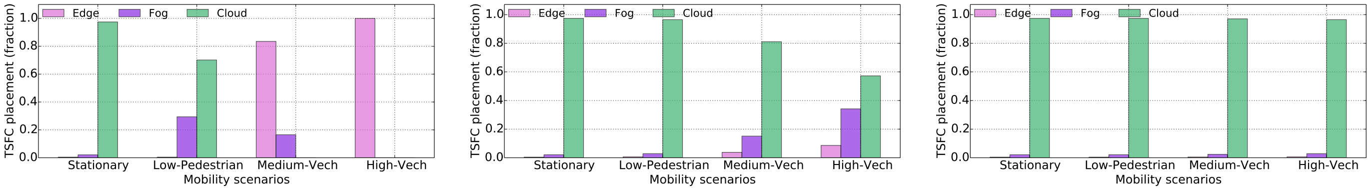


(a) Stationary. (b) Low-pedestrian. (c) Medium vehicular. (d) High vehicular.

Fig. 4. Empirical CDF of controller delay budget for various mobility settings with WiFi 2.4 GHz. Above 99 percentile values are removed for the sake of presentation clarity.



(a) WiFi at 2.4 GHz, small-scale channel changes. (c) WiFi 2.4 GHz, medium-scale channel changes. (e) WiFi at 2.4 GHz, long-term channel changes.



(b) WiFi at 5 GHz, small-scale channel changes. (d) WiFi at 5 GHz, medium-scale channel changes. (f) WiFi at 5 GHz, long-term channel changes.

Fig. 5. Fraction of times each TSFC location is selected as the best choice under the most reactive policy considering: i) small-scale channel changes, i.e., fast fading, (left), ii) medium-scale channel changes, i.e., shadowing, (middle) and iii) long-term channel changes, i.e., handover, (right).

the best location for the TSFC for a particular user speed and flow duration setting using simulations of 10^6 runs. As one-way interface delays, we used the following values: $\tau_f = 50$, $\tau_c = 10$, and $\tau_e = 5$ ms. In each realization, we generate flow duration, user speed, distance from the cell edge, and decorrelation delay values. Using these values, we calculate the best TSFC placement according to our heuristic in (4).

Fig. 4 plots the empiricial CDF of controller delay budget under different mobility scenarios for WiFi APs operating at 2.4 GHz. For stationary case in Fig.4a, all approaches have the same delay budget as the wireless channel remains stable thereby leaving the flow duration T_f as the only limiting factor for controller placement decision. With increasing mobility, the controller delay budget becomes tighter especially for the small scale approach. Next, we show how the delay budget affects the controller placement.

Fig. 5 shows the fraction of times the edge, fog, and cloud-TSFC is chosen as the best TSFC location under 2.4 GHz WiFi network and 5 GHz WiFi network. As expected, for stationary and low-pedestrian scenarios, TSFC logic can be located at the cloud since the wireless channels remain quasi-static and the only limiting factor for the controller's delay budget is the flow duration. The flow duration is longer compared to the delay of interfaces with the cloud-TSFC. Hence, almost all the time cloud-TSFC is preferred for static and low-mobility users according to our heuristic. For some short flows, fog-TSFC would be a better option than the cloud-TSFC for these stationary or low-pedestrian mobility cases as can be seen in Fig. 5a and Fig. 5b. When WiFi operates at 5 GHz, channel changes faster resulting in higher probability of fog-TSFC (Fig.5b) being a better option compared to cloud-TSFC as compared to WiFi at 2.4 GHz (Fig.5a).

On the contrary, for mobile users, best TSFC placement depends on the desired degree of re-activeness. If small-scale changes are to be considered, edge and fog are to be preferred compared to a cloud-TSFC which would be too slow to follow the changes in the wireless links. However, if the network operator prefers to be responsive to only medium or long-scale changes, then cloud-TSFC starts to become a promising choice. Comparing WiFi 2.4 GHz and 5 GHz, there is slight difference between the fraction of times a decision is made. Nevertheless, the trends are the same. If only long-term changes are considered, cloud-TSFC beats almost always the other options.

B. Discussion

Putting the TSFC in the cloud or fog may still be preferred even for short flows. For such short flows, the default rule, i.e. LTE-only bearer, will be applied. That means, small flows may not benefit from multi-RAT operation directly. However, they can still benefit from capacity increase in terms of lower latency due to more spectral resources and lower load on the LTE resources.

We have provided a heuristic to decide on the TSFC location which mainly considers various time scales to react to changes in the network state, e.g., user's channel. However, we have not examined how much acting based on outdated information may degrade the performance. Maintaining the state information of the data forwarding plane at the control plane might be overwhelming. There are some studies, e.g., [17], which investigate the performance implications of acting on imperfect information at the control plane. We leave such an analysis on the impact of outdated information on performance to a future work.

VIII. RELATED WORK

Related work falls into three categories: LWA flow control schemes, link aggregation in multi-RAT networks, and SDN.

LWA flow control schemes: An LWA flow control scheme decides on how to aggregate DL traffic in licensed and unlicensed bands and hence determines the path selection in LWA. Our work is most related to [8] which proposes to split the LWA traffic using assistance from an SDN controller placed at the ePDG node (corresponding to the cloud-controller in our paper) which collects information about the users and their links to several APs. Based on the collected statistics such as link rate, cell load, the SDN controller assigns the AP with the smallest ratio of AP's load to its SNIR to the UE for UEs whose LTE signal is lower than a threshold or if the LTE cell load is above a load threshold. Our paper advocates the control by an SDN controller as [8], but, different from [8], we focus on where the controller should be hosted in the network rather than its control logic. Moreover, in our proposal, actual splitting is always performed at the eNB, in contrast to [8]. For a simple LWA setting, Singh et al.[18] reflect the impact of backhaul latency between eNB and WiFi AP as reduced effective rate for WiFi and use this rate while an eNB decides on the fraction of resources each UE receives in the DL of an eNB. This mentioned work has WiFi only or LWA modes in contrast to LTE only and LWA nodes in conventional LWA. Moreover, it does not clarify when and how frequent the resource assignment should be performed and it remains unclear how multiple WiFi APs shared by several eNBs assigns their resources for the LWA UEs. Using [18]'s water-filling approach, [19] allocates resources in a way to achieve the preset throughput and delay requirements. Lopez-Perez et al. [20] propose to exploit the LWA status reports sent from each UE periodically to the eNB to estimate the WiFi network status, e.g., link capacity, in order to decide on the LWA flow split by selecting the shortest-delay link, LTE or WiFi, for each incoming PDCP PDU. Jin et al.[4] decouple mode selection and bearer scheduling. With the assumption that each AP is connected to only one eNB, first each AP selects a set of UEs among its associated users with the lowest LTE link capacity and therefore should get service also from the WiFi AP. Periodic mode selection takes place every 10 seconds to account for mobile users and the key goal is to achieve intra-cell fairness among users. To minimize the unexpected delays at the WiFi due to competition among different flows, [4] proposes per-bearer queuing at the WiFi. Regarding traffic splitting, the packet is routed through the link with the expected shortest delay. Other related work are [21], [22] which present an LWA prototype.

Link aggregation in multiple radio access technologies (multi-RAT): Flow scheduling in LWA can be considered as a special case of link aggregation in multi-RAT networks. In [23], the flow scheduler optimally determines which fraction of resources of an AP (or BS) each UE is allocated for each of its RATs. For optimal splitting, the scheduler needs all the information about all users (e.g., link rates for each RAT) and therefore it is assumed to be the entry point for all UE traffic. In that respect, the flow scheduler in [23] must be in the LTE core network and it hosts both TSF and TSFC functionality so that it has a global view of the network. Amount of traffic routed through each RAT is proportional to the total promised throughput at respective RAT, e.g., if WiFi has 30 Mbps whereas LTE has 15 Mbps capacity for this user, ratio of traffic on WiFi to LTE RAT is 30/15. While [23] provides a closed form solution, it is not clear how latency to retrieve network state information would affect the operation of the flow scheduler and how often the flow scheduler should solve the traffic splitting problem.

Software Defined Networks (SDN): Our work is inspired by the wide use of SDN principles of separating the data and control plane in areas ranging from cloud-based solutions for spectrum assignment in Cognitive Radio networks [24], control for channel assignment and airtime management of residential WiFi networks [25] and dynamic slicing of fronthaul and backhaul data networks [26], [27] to edge-based solutions for network load balancing in residential WiFi networks [28].

IX. CONCLUSIONS

In this work, we have presented an alternative LWA architecture which has merits of flexibility at the control plane and also the simplicity at the data forwarding plane. Taking the advantage of the introduced flexibility, we discussed how the traffic split function logic can be moved to different parts of the network and how it might affect the performance of an LWA network. As future work, we plan to implement a remote TSFC using current SDN tools and analyze the feasibility and performance in practise. To assess how the location of TSFC affects the network performance, we plan to provide a through performance analysis using system-level simulations in NS-3 on throughput and delay achieved under the three cases: edge-TSFC, fog-TSFC, and cloud-TSFC. Such an analysis can also include analysis on the centralization gain under reactive and proactive TSFCs, impact of delayed/outdated control plane information, and also the scalability of each TSFC architecture.

X. ACKNOWLEDGMENT

This work received funding from the European Horizon 2020 Programme in the scope of the EU eWINE Project (Nr. 688116).

REFERENCES

[1] P. Nuggehalli, "LTE-WLAN aggregation [industry perspectives]," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 4–6, 2016.

[2] Qualcomm, "LTE-U Technology and Coexistence," https://goo.gl/tZV9VG, May 2015, online; accessed 29 Aug. 2017.

[3] S. Sirotkin, "LTE-WLAN Aggregation(LWA): Benefits and Deployment Considerations," Intel Corporation, Tech. Rep., 2016.

[4] B. Jin, S. Kim *et al.*, "Aggregating LTE and Wi-Fi: Towards intra-cell fairness and high TCP performance," *IEEE Transactions on Wireless Communications*, 2017.

[5] Y. Chen, M. Ding *et al.*, "Dynamic reuse of unlicensed spectrum: An inter-working of LTE and WiFi," *IEEE Wireless Communications*, vol. 24, no. 5, pp. 52–59, 2017.

[6] C. Rosa, K. Pedersen *et al.*, "Dual connectivity for LTE small cell evolution: Functionality and performance aspects," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 137–143, 2016.

[7] B. Heller, R. Sherwood, and N. McKeown, "The controller placement problem," in *Proceedings of the first workshop on Hot topics in software defined networks*. ACM, 2012, pp. 7–12.

[8] S. Anbalagan, D. Kumar *et al.*, "SDN-assisted efficient LTE-WiFi aggregation in next generation IoT networks," *Future Generation Computer Systems*, 2017.

[9] H.-L. Määttanen, G. Masini *et al.*, "LTE-WLAN aggregation (LWA) in 3GPP Release 13 & Release 14," in *IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2017, pp. 220–6.

[10] R. Burbidge, "Liaison from 3GPP on LWA and LWIP," March 2016.

[11] Open Networking Foundation (ONF), "OpenFlow switch specification version 1.5.1 (protocol version 0x06)," Tech. Rep., March 2015.

[12] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[13] S. Bayhan and A. Zubow, "Optimal mapping of stations to access points in enterprise wireless local area networks," in *Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems*. ACM, 2017, pp. 9–18.

[14] J. Yang, W. Li *et al.*, "Characterizing and modeling of large-scale traffic in mobile network," in *IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2015, pp. 801–806.

[15] M. Maternia, S. E. El Ayoubi *et al.*, "5G PPP use cases and performance evaluation models," *5G PPP*, vol. 1, 2016.

[16] T. Abbas, K. Sjöberg, J. Karedal, and F. Tufvesson, "A measurement based shadow fading model for vehicle-to-vehicle network simulations," *International Journal of Antennas and Propagation*, vol. 2015, 2015.

[17] B. Azimdoost, C. Westphal, and H. R. Sadjadpour, "Resolution-based content discovery in network of caches: Is the control traffic an issue?" *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2943–2955, July 2017.

[18] S. Singh, M. Geraseminko *et al.*, "Proportional fair traffic splitting and aggregation in heterogeneous wireless networks," *IEEE Communications Letters*, vol. 20, no. 5, pp. 1010–13, 2016.

[19] G. Smpokos, A. Lioumpas *et al.*, "Performance aware resource allocation and traffic aggregation for user slices in wireless HetNets," in *IEEE 22nd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2017, pp. 1–5.

[20] D. López-Pérez, D. Laselva *et al.*, "Long term evolution-wireless local area network aggregation flow control," *IEEE Access*, vol. 4, pp. 9860–9869, 2016.

[21] Y.-J. Shih, "Control plane design and implementation for LTE-WLAN aggregation (LWA)," *ICT Journal*, no. 168, pp. 78–85, 2016, available at https://ictjournal.itri.org.tw.

[22] Y.-B. Lin, Y.-J. Shih, and P.-W. Chao, "Design and implementation of LTE RRM with switched LWA policies," *IEEE Transactions on Vehicular Technology*, 2017.

[23] S. Singh, S.-p. Yeh, N. Himayat, and S. Talwar, "Optimal traffic aggregation in multi-RAT heterogeneous wireless networks," in *IEEE Int. Conf. on Communications Workshops (ICC)*, 2016, pp. 626–631.

[24] A. Zubow, M. Döring, M. Chwalisz, and A. Wolisz, "A sdn approach to spectrum brokerage in infrastructure-based cognitive radio networks," in *Dynamic Spectrum Access Networks (DySPAN), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 375–384.

[25] A. Patro and S. Banerjee, "Coap: A software-defined approach for home wlan management through an open api," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 18, no. 3, pp. 32–40, 2015.

[26] E. Garcia-Villegas, D. Sesto-Castilla *et al.*, "Senseful: an sdn-based joint access and backhaul coordination for dense wi-fi small cells."

[27] M. Vincenzi, A. Antonopoulos *et al.*, "Multi-tenant slicing for spectrum management on the road to 5G," *IEEE Wireless Communications*, vol. 24, no. 5, pp. 118–125, October 2017.

[28] P. Gawlowicz, S. Zehl, A. Zubow, and A. Wolisz, "Nxwlan: Towards transparent and secure usage of neighbors," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2017, pp. 1–10.