This paper was accepted for publication at the 2018 IEEE International Workshop on Computer-Aided Modeling Analysis and Design of Communication Links and Networks (CAMAD), September 2018, Barcelona, Spain.

End-to-End Simulation of Integrated Access and Backhaul at mmWaves

Michele Polese°, Marco Giordani°, Arnab Roy†, Sanjay Goyal†, Douglas Castor†, Michele Zorzi°
email:{polesemi,giordani,zorzi}@dei.unipd.it, {name.surname}@interdigital.com
°Consorzio Futuro in Ricerca (CFR) and University of Padova, Italy
†InterDigital Communications, Inc., USA

***Abstract*—Recently, the millimeter wave (mmWave) bands have been investigated as a means to support the foreseen extreme data rate demands of next-generation cellular networks (5G). However, in order to overcome the severe isotropic path loss and the harsh propagation experienced at such high frequencies, a dense base station deployment is required, which may be infeasible because of the unavailability of fiber drops to provide wired backhauling. To address this challenge, the 3GPP is investigating the concept of Integrated Access and Backhaul (IAB), i.e., the possibility of providing wireless backhaul to the mobile terminals. In this paper, we (i) extend the capabilities of the existing mmWave module for ns-3 to support advanced IAB functionalities, and (ii) evaluate the end-to-end performance of the IAB architecture through system-level full-stack simulations in terms of experienced throughput and communication latency. We finally provide guidelines on how to design optimal wireless backhaul solutions in the presence of resource-constrained and traffic-congested mmWave scenarios.**

***Index Terms*—5G, millimeter wave, Integrated Access and Backhaul, 3GPP, NR.**

I. Introduction

The 3rd Generation Partnership Project (3GPP) has recently started investigating new Radio Access Technologies (RATs) as enablers for the performance requirements of next-generation wireless systems, e.g., the so-called 5th generation (5G) network architecture [1]. 5G will be characterized by very stringent requirements in terms of latency, jitter and reliability, and is expected to provide the users with unprecedented data rates (e.g., on the order of gigabits per second [2]), in line with the mobile data traffic predictions for 2020 and beyond [3]. In this context, the millimeter wave (mmWave) bands above 10 GHz will play a key role, thanks to the very large bandwidths available at such high frequencies (up to 400 MHz per carrier, according to the latest 3GPP NR specifications [4]). Moreover, the small size of antennas at mmWaves makes it practical to build very large antenna arrays and obtain high beamforming gains, thereby overcoming the high propagation loss at such high frequencies and increasing the spectral efficiency. On the other hand, mmWave signals do not penetrate most solid materials and are subject to high signal attenuation and reflection, thus limiting the communication range of the mmWave infrastructures. Although such an effect can be partially alleviated by configuring directional transmissions, a tracking mechanism is required to maintain the alignment of the transmitter and receiver beams, an operation that may increase the system overhead and lead to throughput degradation [5].

The combination of the high propagation loss and the blockage phenomenon calls for a high-density deployment of Next Generation Node Bases (gNBs) (i.e., the base stations in NR terminology), to guarantee Line-of-Sight (LOS) links at any given time and decrease the outage probability [6]. In such a deployment, providing wired backhaul to each of the gNBs will be inevitably costly for network operators and, as a result, more economically sustainable solutions have been recently investigated by the 3GPP as part of the Integrated Access and Backhaul (IAB) Study Item (SI) [7] for NR. The idea is to have wireless backhaul [8], i.e., a fraction of gNBs with traditional fiber-like backhaul capabilities and the rest of the gNBs connected to the fiber infrastructures wirelessly, possibly through multiple hops.

The performance of this deployment paradigm is typically assessed in terms of hop count and bottleneck Signal-to-Noise-Ratio (SNR), e.g., in [9], [10], which, however, prevents the comprehensive evaluation of the end-to-end performance of the different backhaul approaches. In this regard, discrete-event network simulators enable full-stack simulation of complex and realistic scenarios, and therefore represent a viable tool for the setup of more accurate system-level analysis.

In this paper, we therefore present the extension with the IAB features of the mmWave module for ns-3, which already models the mmWave channel and the Physical (PHY) and Medium Access Control (MAC) layers of the mmWave protocol stack [11]. This extension can support both single- and multi-hop deployments and autonomous network configuration, and features a detailed 3GPP-like protocol stack implementation. Moreover, new scheduling mechanisms have been developed in order to support the sharing of access and backhaul resources.

We also evaluate the performance of the proposed IAB architecture in an end-to-end environment in terms of experienced throughput and latency, considering realistic traffic models. Our preliminary results demonstrate that the configuration of a wireless backhaul deployment has the potential to increase the overall network throughput as well as to reduce the communication latency in case of congested networks, and therefore represents a practical solution in future mmWave networks. Our simulator can be used for the design of enhanced multi-hop routing strategies and scheduling algorithms, and for the determination of the best strategy for the deployment of a wireless backhaul solution.

The rest of the paper is organized as follows. Sec. II describes the characteristics of the 3GPP SI on IAB and the potential of this solution for NR deployments. Sec. III presents the implementation of the IAB features in ns-3, while Sec. IV provides a preliminary evaluation of the end-to-end performance of the IAB nodes in a realistic mmWave scenario. Finally, Sec. V concludes the paper and discusses possible extensions of this work.

II. STATE OF THE ART ON INTEGRATED ACCESS AND BACKHAUL

Research on wireless backhaul solutions has been carried out in the past at frequencies below 6 GHz, e.g., in the WLAN domain [12] and as part of the Long Term Evolution (LTE) standardization activity with a single wireless backhaul hop [13]. The practical implementation of wireless multi-hop networks, however, never really turned into a commercial deployment due to practical limitations including, but not limited to, (i) scalability issues [8], (ii) the scheduling constraints between hops [14], and (iii) the large overhead for maintaining multi-hop routes [15].

Nonetheless, with the recent advancements in mmWave communication and leveraging highly directional beamforming, the integration of wireless backhaul and radio access is being considered as a promising solution for 5G cellular networks. In [16], the authors demonstrated that the noise-limited nature of large-bandwidth mmWave networks offer interference isolation, thereby providing an opportunity to incorporate self-backhauling in a mesh small-cell deployment without significant throughput degradation. Paper [17] showed that wireless backhaul over mmWave links can meet the expected increase in mobile traffic demands, while paper [18] evaluates the energy efficiency of mmWave backhaul at different frequencies. The authors in [19] further evaluated the performance of the integration between access and backhaul and determined the maximum total network load that can be supported using the IAB architecture.

Along these lines, the 3GPP is also focusing on IAB for 3GPP NR [7], to design an advanced wireless relay which overcomes the limitations of the traditional LTE implementation and makes it possible to flexibly deploy self-backhauled NR base stations. According to [7], NR cellular networks with IAB functionalities will be characterized by (i) the possibility of using the mmWave spectrum; (ii) the integration of the access and backhaul technologies, i.e., using the same spectral resources and infrastructures to serve both mobile terminals in access as well as the NR gNBs in backhaul [20]; (iii) a higher flexibility in terms of network deployment and configuration with respect to LTE, i.e., the possibility of deploying plug-and-play IAB nodes capable of self-configuring and self-optimizing themselves [21]. According to [7], [21], 5G IAB relays will be used in both outdoor and indoor scenarios, possibly with multiple wireless hops, with the final goal of extending the coverage of cell-edge users, avoiding service

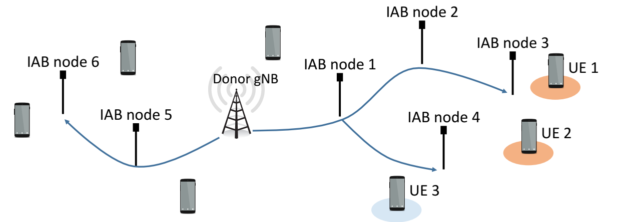

Fig. 1: Example of IAB architecture, with a single donor and multiple downstream IAB nodes.

unavailability, and increasing the efficiency of the resource allocation. Both in-band and out-of-band backhaul will be considered, with the first being a natural candidate for a tighter integration between access and backhaul.

Despite its clear strengths, the design of IAB solutions in mmWave systems is a research challenge that is still largely unexplored. Most of the existing literature does not consider a channel characterized by the full channel matrix, nor realistic beamforming patterns. Moreover, the prior art lacks considerations on the end-to-end performance of the self-backhauling architectures, which are in turn part of our original contributions.

III. IAB IN NS-3 MMWAVE

The ns-3 mmWave module, described in [11], enables the simulation of end-to-end cellular networks at mmWave frequencies. It features a complete stack for User Equipments (UEs) and gNBs, with a custom PHY layer, described in [22], the 3GPP mmWave channel model and, thanks to the integration with ns-3, a complete implementation of the TCP/IP protocol stack.

As mentioned in the previous sections, IAB will be important for NR ultra-dense mmWave deployments.[1] Therefore, in order to increase the realism and the modeling capabilities of the ns-3 mmWave module, we implemented an IAB framework that will be described in the following sections. It features a new ns-3 `NetDevice`, the `MmWaveIabNetDevice` with a dual stack for access and backhaul, an extension of the ns-3 mmWave module schedulers, and network procedures to support IAB nodes in a simulation scenario. Moreover, we simulate the wireless relaying of both data and control plane messages, in order to accurately model the IAB operations.

An example of IAB network that can be now supported by ns-3 is shown in Fig. 1. In particular, we consider a tree architecture, with the root being a donor gNB, i.e., a base station with a wired connection to the core network. Therefore, this is not a traditional mesh architecture, which is used, for example, for random-access-based backhaul technologies such as IEEE 802.11 [23], in which there is no strict parent/child relationship between network nodes. In a cellular context, it is necessary to define a tree structure because every

[1]The 3GPP SI on IAB is still ongoing and is scheduled for completion as part of Release 16. We therefore do not preclude in the future to further extend the features of the ns-3 IAB module to make it fully compliant with the latests 3GPP specifications on this topic.

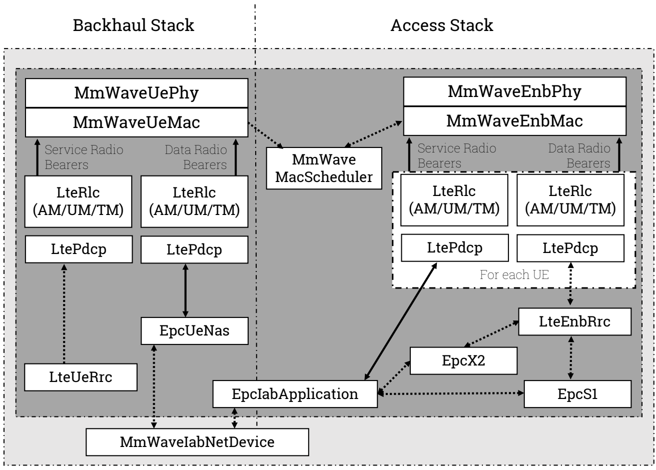

Fig. 2: Protocol stack and organization of the ns-3 classes for an IAB node.

communication is scheduled [1], i.e., the base station assigns specific time and frequency resources for downlink or uplink communication with any connected UE. Therefore, given that the access and the backhaul share the same resources, then also communication between the gNB and any IAB node must be scheduled. Notice that the connection between a parent and a child node can change with handover procedures[2], for example if the link quality between them degrades because of blockage.

In the following paragraphs we will describe the protocol stack that is deployed in each IAB node, the scheduling mechanism, and how to set up a simulation with IAB features.

A. IAB node

As mentioned in [7], the IAB nodes should re-use the specifications for the access stack of NR as much as possible. At the moment, there are a few protocol stacks being discussed in the 3GPP [24]. All of them, however, include PHY, MAC and Radio Link Control (RLC) layers, and differ because of the support of layer-2 (i.e., RLC or Packet Data Convergence Protocol (PDCP)) or layer-3 relaying. Given the need for a flexible solution, able to adapt to the direction that the 3GPP will take, we implemented a light layer-3 relaying solution, i.e., each backhaul radio bearer is set up locally, and a middle layer above the PDCP handles the forwarding of the packets from the access to the backhaul PDCPs. Fig. 2 shows the protocol stack for an IAB node and the classes that model it.

The main novelties are the `MmWaveIabNetDevice` and the `EpcIabApplication` classes. The first is an extension of the ns-3 `NetDevice` class, and, similarly to the `NetDevice` implementations of the UE and gNB, holds pointers to all the objects that model the other layers of the protocol stack. Moreover, it is internally used in the ns-3 model to forward packets between an instance of the `EpcUeNas` class in the backhaul stack and the `EpcIabApplication` in the access stack.

[2]We will introduce support for this functionality in the next iteration of the module.

The `EpcIabApplication`, instead, implements the main logic related to the control and data plane management in the IAB node. In particular, for the data plane, the `EpcIabApplication` class is in charge of applying the forwarding rules for local UEs, i.e., those directly connected to the IAB node this class belongs to, and for remote UEs, i.e., those connected to downstream IAB nodes. In this case, the traffic will be forwarded to the local bearer mapped to the downstream IAB device. More details on how the routing is performed will be given in Sec. III-B. This class is also responsible for the processing and forwarding of control packets for the interfaces toward the core network and the other neighboring gNBs. When a control message is received on either the access or the backhaul interface, the `EpcIabApplication` checks if it is a local message, i.e., if the destination is the Radio Resource Control (RRC) layer of the current IAB node, and, if this is the case, forwards the packet to the RRC. Otherwise, as done in the data plane, the packets are relayed via one of the downstream IAB nodes.

The other classes are the same as those used in the UE protocol stack (for the backhaul) and gNB protocol stack (for the access). The consequence is that, in the access, the UEs in the scenario consider the IAB node as a normal gNB, and, similarly, in the backhaul, the parent gNBs and/or IAB nodes consider the IAB child as a UE. Therefore, there is no need to adapt the UE and gNB ns-3 implementations to support the IAB feature. The only change is the extension of the gNB schedulers, to support the multiplexing of access and backhaul in the same resources, and the introduction of a new interface between the access and backhaul MAC layers. These extensions will be described in Sec. III-C. Nonetheless, additional enhancements can be introduced in future releases, to improve the overall performance of the IAB protocol stack and track the 3GPP SI and specifications on IAB.

B. Single- and multi-hop control procedures

Given that the 3GPP is still considering IAB as an SI, there are no standard specifications yet on control procedures to support IAB networks. Nonetheless, the SI [7] specifies that both single- and multi-hop topologies should be considered, and that the IAB node should be able to autonomously connect to the network, adapt the access and backhaul resource partitioning and, eventually, independently update the parent node in case of blockage. All these features require specific control procedures, and, given the high level of detail of the ns-3 model, we implemented a number of realistic control procedures, which involve an exchange of messages on the wireless backhaul links to set up and automatically configure the network. These can be easily updated to implement different network procedures that the 3GPP may specify in the future.

In particular, we assume that the parent IAB node for a backhaul link terminates the NG control interface to the core network (i.e., the NR equivalent of the LTE S1 interface) [25], and that it takes care of forwarding the control messages towards the network servers that host the Access and Mobility

Management Function (AMF). Moreover, the IAB node has a similar role with respect to the UEs connected to it, as would happen with a traditional wired gNB. Thanks to this design, the differences with respect to the 3GPP specifications for the access stack are minimized. This configuration makes it possible to seamlessly support both single- and multi-hop deployments, given that the architecture of the upstream portion of the network is transparent to each IAB node, which will simply relay all of its packets to the parent. Furthermore, for the purpose of packet transport in the backhaul network, we exploit GPRS Tunneling Protocol (GTP) tunnels from each IAB node to the relevant element in the core network (i.e., the server with control functions or the packet gateway). Each data bearer of all the UEs (and IAB nodes, for the backhaul part) is associated with a unique tunneling ID, and all the packets sent on backhaul links will be associated with a GTP header carrying that ID.

We also implemented realistic autonomous access and configuration procedures for the IAB nodes. When the IAB selects its parent node during the Initial Access (IA) procedure[3], the parent sends an initial message to the AMF, which will reply with the configuration for the backhaul bearer between the IAB node and its parent. These messages will be relayed by all the IAB nodes in the path between the parent and the donor gNB, and each of them will register the presence of an additional downstream IAB device. Notice that there may be multiple IAB children for each parent, therefore the parent has to match the new downstream node to the correct child to correctly route the other control and data packets.

For the UEs, there is no difference between a wireless relay and a gNB with a wired connection to the core network. Therefore, the UE's IA procedure does not change, and the IAB node will take care of forwarding the relevant control messages to the AMF and the other network functions involved in the IA. Moreover, the upstream relays and the donor gNB will exploit the control messages for the UE's IA to associate to each IAB bearer the total number of downstream UEs. For example, by considering the deployment in Fig. 1, if UEs 1 and 2 connect to IAB node 3, and UE 3 connects to IAB node 4, then IAB node 1 will know that the backhaul bearer towards IAB node 2 will carry the traffic for 2 UEs, and that towards node 4 will account for a single UE. This information could be exploited by advanced IAB MAC schedulers. Finally, during the UE IA procedure, each gNB associates the GTP tunneling ID of the bearers of downstream UEs to a local IAB child, so that, when a backhaul packet is received, the gNB uses the information in the GTP header to correctly route the packet.

C. Backhaul-aware dynamic scheduler

The MAC and the associated scheduler are a key component in the design of scheduled wireless relay architectures in which the resources between the access and the backhaul are shared. In order to avoid self-interference between access and backhaul, indeed, there is a need to multiplex the two interfaces. In our implementation, we consider Time Division Multiple Access (TDMA), but we plan to extend the support to spatial division multiplexing in future releases, to harness the directionality of mmWave communications. Moreover, the scheduler is usually not part of the 3GPP specifications, and, therefore, equipment vendors have the possibility of designing custom solutions in this domain.

We opted for a distributed scheduling solution, in order to minimize the difference in the scheduling mechanism with respect to a traditional access-only scenario, and to limit the amount of control overhead that a centralized solution would require. Therefore, in the ns-3 mmWave IAB module, each gNB (either wired or wireless) schedules the resources for its access interface (i.e., for both UEs and IAB children) independently of the other gNBs, as would happen in a traditional network without IAB. In a TDMA setup, however, the IAB node cannot schedule resources in the time and frequency slots already allocated to the backhaul by their parent. Therefore, if at time t the relay has to perform a scheduling decision for subframe $t + \eta$, then it has to be already aware of the scheduling decision of its parent for $t + \eta$. Given a delay ϵ for the communication of scheduling information between the parent and the current relay, then the parent should perform its scheduling decisions for $t + \eta$ at time $t - \epsilon$.

In order to efficiently address this issue, we implemented a *look-ahead backhaul-aware scheduling* mechanism. The backhaul-aware component is given by a new interface between the access and the backhaul MAC layers. The backhaul MAC layer is seen as a UE by the parent node, and thus will receive Downlink Control Informations (DCIs) with the scheduling and modulation and coding scheme information for η subframes in advance. Then, the backhaul MAC shares DCI with the scheduler of the IAB node (in the access stack), which registers the resources occupied by backhaul transmissions for the relevant subframe.

The look-ahead mechanism, additionally, makes it possible to adjust the value of η according to the maximum number of downstream relaying hops N from the current gNB to the farthest IAB node: the gNB schedules ahead by $\eta = N + 1$ subframes[4], and propagates this information with a DCI to the UEs and IAB nodes connected to it. In turn, these IAB nodes will schedule ahead by at most $\eta = N$ subframes. Each of them will consider the time and frequency resources allocated for their downlink or uplink backhaul transmission as busy, and will schedule access resources for their UEs and, eventually, for IAB nodes in unallocated resources. For example, by considering Fig. 1, the maximum number of hops from the donor gNB is 3. Therefore, the donor will schedule ahead by 4 subframes. On the other hand, IAB node 2 has a single hop to the farthest relay, thus it will schedule ahead by 2 subframes. Notice that, thanks to the procedures introduced in

[3]Initial access is the procedure by which a mobile terminal establishes an initial physical link connection with a cell, a necessary step to access the network. For a complete overview of the most relevant works on IA for 3GPP NR scenario we refer the reader to [26], [27].

[4]The additional subframe with respect to N is needed because the farthest IAB node (without IAB children) has to schedule its resources at least one subframe in advance, in order to transmit the DCI beforehand to its UEs

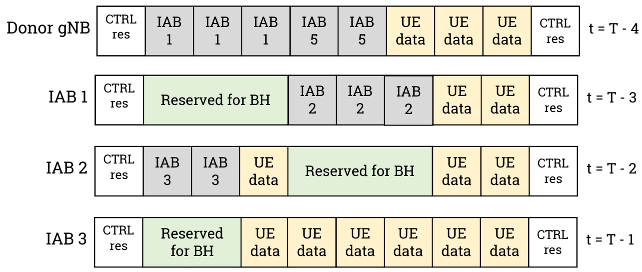

Fig. 3: Example of resource allocation for time T with a look-ahead scheduler at the donor gNB and IAB nodes 1, 2 and 3 in the deployment of Fig. 1.

Sec. III-B, there is no need to manually tune the η parameter, which is automatically configured according to the structure of the IAB tree, and can be updated in case of variations in the architecture of the network.

We added the look-ahead and backhaul-aware capabilities to two of the ns-3 mmWave module schedulers, i.e., the `MmWaveFlexTtiMacScheduler` class, which models a Round Robin (RR) scheduler, and the `MmWaveFlexTtiPfMacScheduler` class, which implements a Proportional Fair (PF) scheduling algorithm. Moreover, in a TDMA setup, with shared resources between the access and the backhaul, it is important to make sure that the parent gNB does not schedule all of the available resources to a single IAB node (e.g., if it is the only active terminal connected to the parent). Otherwise, the child IAB node would not be able to allocate any resource to the access. Therefore, we limit the maximum number of time and frequency resources that can be allocated to an IAB device to half of the total available resources.

An example of resource allocation is shown in Fig. 3, where a total number of 10 time and frequency resources are dynamically allocated to access and backhaul links. In particular, we refer to the deployment in Fig. 1, and present a possible resource partitioning for the donor gNB, IAB nodes 1, 2 and 3 and the UEs connected to these gNBs. As can be seen, each IAB node does not allocate access transmission in the resources reserved for its backhaul, but can exploit all of the other resources for communication with other relays and the UEs, including those allocated by one of the upstream nodes to other backhaul links. While in general this may increase the interference, it must be noticed that, at mmWave frequencies, the large antenna arrays that can be built and the resulting directional transmissions that can be established have the potential to provide increased spatial reuse and isolation, thereby guaranteeing reduced interference [28]. Moreover, interference-aware schedulers can be designed and tested with the simulator. Finally, it is possible to update the allocation on-the-fly, to dynamically adapt to changed channel conditions and traffic requirements from the different connected terminals.

Fig. 3, however, also highlights one of the main bottlenecks of an IAB architecture, i.e., the fact that the donor gNB needs to serve not only its own users, but also all the downstream relays, carrying traffic from many other UEs. On one hand, the amount of data that can be exchanged on a backhaul link in each time and frequency resource is generally higher than the equivalent for a gNB-UE link, thus the backhaul will probably require fewer resources. Indeed, the backhaul link between two gNBs has usually a better quality than that between a gNB and a UE, given that a backhaul link is expected to be in LOS, and that a larger number of antennas can be deployed in a relay than in a UE. On the other hand, the scalability of an IAB deployment has an intrinsic limitation due to the resource sharing between the access and the backhaul link. Therefore, efficient scheduling algorithms will be key for high-performance IAB networks. This makes the ns-3 mmWave module with the IAB integration a valuable platform for researchers interested in IAB networks, given that it offers a lean interface to the scheduler implementations, which can be easily extended to test new IAB scheduling paradigms in realistic end-to-end scenarios.

D. Simulation setup

The setup of a simulation with the IAB feature resembles that of a simulation with traditional wired-only backhaul. An extensive description of how to configure an ns-3 simulation script for the mmWave module is provided in the tutorial in [11]. We added two auxiliary methods in the `MmWaveHelper` class, which hides from the ns-3 user much of the complexity related to the configuration of the mmWave Radio Access Network (RAN) and core network. Similarly to the methods used to set up UEs and gNBs, the `InstallIabDevice` method returns a `NetDevice` properly configured, with the stack described in Fig. 2.

The initial attachment of each IAB node to its parent gNB is performed by the methods `AttachIabToClosestWiredEnb` or `AttachIabToBestNodeHQF`. The latter scans the signal quality of the available IAB nodes or wired donors, and selects that with the highest SNR. Moreover, it avoids the creation of loops in the network tree. These helper methods, moreover, automatically register the new IAB nodes to the control entities in the core network, and define the default radio bearer that will be used for the backhaul link. Finally, by default, the UEs in the ns-3 mmWave module perform the initial attachment as soon as the simulation starts, i.e., at simulation time $t_s = 0$. Therefore, we added the `AttachToClosestEnbWithDelay` method that delays by D seconds the initial attachment of UEs to the chosen gNBs, either wired or wireless. This method can be used to let the UEs perform IA only after the IAB nodes have completed their IA and backhaul bearer setup.

IV. Example results

In this section, we validate the implementation of the IAB features for the ns-3 mmWave module through simulations. We illustrate some preliminary results related to the coverage performance of an IAB deployment in a mmWave environment. The considered scenario is a Manhattan grid, with blocks of

Parameter	Value
mmWave carrier frequency	28 GHz
mmWave bandwidth	1 GHz
3GPP Channel Scenario	Urban Micro
mmWave max PHY rate	3.2 Gbit/s
MAC scheduler	Round Robin
Subframe duration	1 ms
Donor gNB to remote server latency	11 ms
RLC buffer size B_{RLC} for UEs	10 MB
RLC buffer size B_{RLC} for IAB nodes	40 MB
RLC AM reordering timer	2 ms
UDP rate R	$\{28, 224\}$ Mbit/s
UDP packet size	1400 byte
Number of independent simulation runs	50

TABLE I: Simulation parameters

50 m for each side, and with 10 m between each block, for a total area of 0.053 km^2. A gNB with a wired connection to the core network is placed at the center of the scenario, while the number of IAB nodes (i.e., gNBs with wireless backhaul functionalities) varies from 0 to 4. The IAB nodes are one block in each direction away from the donor (i.e., at a distance of 85 m), and they are in LOS (e.g., placed on the building rooftops). Each relay directly connects to the wired donor wirelessly, thus this scenario only considers single-hop transmissions.[5] 40 users are randomly placed outdoors using the new ns-3 `OutdoorPositionAllocator` method, and connect to the closest gNB, either wired or wireless. Each UE downloads content from a remote server at a constant rate $R = \{28, 224\}$ Mbit/s using UDP as the transport protocol. These two different source rates are used to test the network in different congestion conditions. Finally, the MAC layer performs Hybrid Automatic Repeat reQuest (HARQ) retransmissions, and the RLC layer uses the Acknowledged Mode (AM) to provide additional reliability. The scheduler is Round Robin, with the look-ahead backhaul-aware mechanisms described in Sec. III-C. The other simulation parameters are in Table I.

We consider two different end-to-end metrics, i.e., the experienced throughput and the application-layer latency averaged over multiple independent runs. Fig. 4 investigates three different throughput values for different source rates R and varying the number of IAB relays. We observe that, for the low source rate scenario (i.e., $R = 28$ Mbit/s), the total throughput remains almost constant, while, in the congested scenario (i.e., $R = 224$ Mbit/s) the rate progressively increases with the number of relays. This shows that, in the considered Manhattan scenario, the relays extend the area in which the mobile terminals can benefit from the coverage of their serving infrastructures and, in particular, have the potential to improve the quality of the access link between the cell-edge users and the donor gNB, thereby guaranteeing higher capacity.

[5]Although our simulator enables multi-hop relaying operations, for the tractability of the simulation in this paper we only focus on single-hop transmissions, and we leave the analysis of the multi-hop architecture as part of our future work.

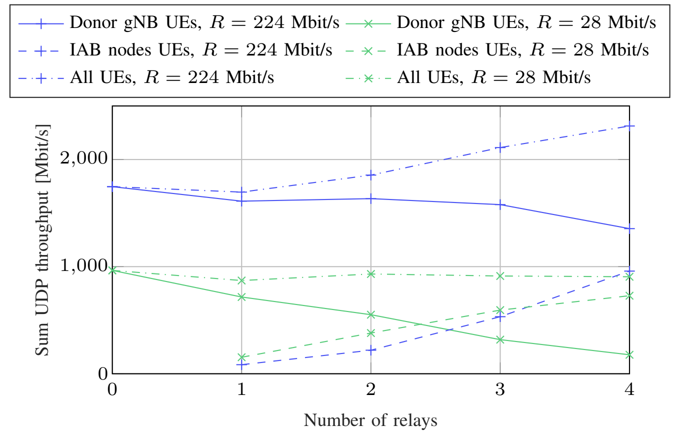

Fig. 4: Sum end-to-end throughput for different source rate R and number of relays. The total throughput is the sum of the throughput of all the users, while the wired (or IAB nodes) sum throughput refers to the aggregate throughput of UEs connected to the donor (or the relays, respectively).

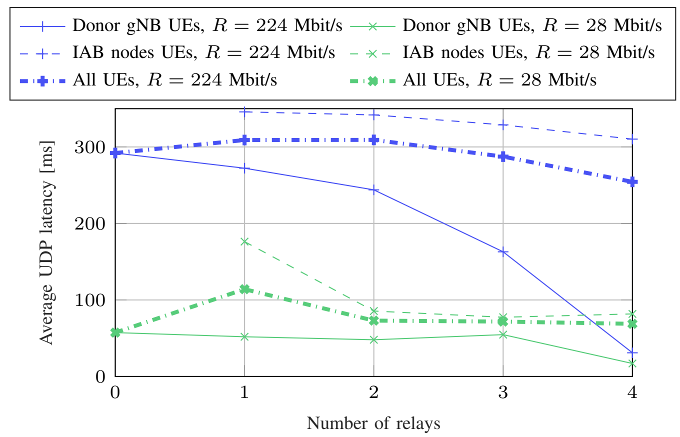

Fig. 5: Average end-to-end latency for different source rate R and number of relays. We report the average latency considering all the UEs, or only those connected to the wired gNB or wireless relays. The dotted black line represents the average latency of the configuration with 0 relays.

The average latency is shown in Fig. 5. We see that, in a Manhattan grid scenario, the average latency of the UEs directly connected to the wired gNB decreases as a result of increasing the number of wireless relays. Indeed, if the relays are used, the wired gNB will serve fewer users, i.e., those with the best channel quality, and will avoid allocating resources to cell-edge users which, generally, require a high number of HARQ and RLC retransmissions. Although these benefits are particularly evident in the $R = 224$ Mbit/s case, a latency improvement is also observed for the non-congested scenario (i.e., $R = 28$ Mbit/s) when four relays are deployed.

On the other hand, from Fig. 5 we notice that the average latency of the users attached to IAB nodes increases with respect to the configuration without relays, especially when just one or two wireless relays are deployed. This is mainly due to the buffering that occurs in the backhaul. In an IAB context, indeed, the backhaul and access resources are shared, thus the IAB nodes and the UEs attached to the donor contend for the same resources. With an RR scheduler, a similar number of transmission opportunities is allocated to the IAB nodes and to the UEs, but the relays generally have more data to transmit than each single UEs. Consequently, the buffering latency at the RLC layer of the relays increases. Nonetheless, for the congested scenario (i.e., $R = 224$ Mbit/s), the overall average latency when more than three relays are deployed (i.e., 287 and 250 ms for three and four relays, respectively) is equivalent or lower than that in the configuration with the donor gNB only (i.e., 292 ms), as shown in Fig. 5.

The above discussion exemplifies how an IAB architecture introduces both opportunities and challenges. From one side, the deployment of wireless relays is a viable approach to increase the coverage of cell-edge users, i.e., the most resource-constrained network entities, thereby promoting fairness in the whole network. Moreover, the presence of the wireless backhaul nodes has the potential to reduce the communication latency in case of congested networks. From the other side, the IAB nodes may degrade the throughput and latency performance of some UEs, i.e., those with the best channel quality, whose traffic would have been successfully handled even in traditional wired backhaul scenarios. It becomes therefore fundamental to determine the optimal number of wireless backhaul nodes to be deployed and to design efficient scheduling policies, according to the context and considering the constraints imposed by the available network and economic resources. This research challenge will be part of our future work.

V. Conclusions and Future Work

The integration between access and backhaul will likely play a key role in the next generation of wireless networks operating in the mmWave band. In this paper, after reviewing prior work on self-backhauling and the potential of this solution for 3GPP NR deployments, we presented the first implementation of IAB for the ns-3 mmWave module[6]. The simulator, which features the 3GPP mmWave channel model and a complete characterization of the TCP/IP protocol stack, now also implements the wireless relaying operations on both the data and the control planes, thereby accurately modeling the operations of an IAB network. We believe that this tool can be used by researchers to understand the main limitations and the performance gains that IAB networks can provide, and to evaluate new integrated scheduling algorithms and multi-hop routing strategies with a realistic, end-to-end protocol stack.

We have also provided the first preliminary end-to-end performance evaluation, in terms of experienced data rate and communication latency, of mmWave nodes in an IAB scenario. We showed that the IAB architecture may represent a viable solution to efficiently relay the traffic of cell-edge users in very congested networks.

This work opens up some particularly interesting research directions. More specifically, we plan to investigate how to design advanced backhaul path selection policies as well as to determine the best degree of migration from a fully-wired backhaul deployment to a wireless backhaul solution when considering both economic and performance trade-offs. Moreover, we will further extend the ns-3 mmWave module with additional IAB features, in order to address mobility scenarios, and keep track of the 3GPP specifications on this topic.

[6]The code can be found at https://github.com/signetlabdei/ns3-mmwave-iab.

References

[1] 3GPP, "NR and NG-RAN Overall Description," TS 38.300 (Rel. 15), 2018.

[2] M. Shafi, A. F. Molisch, P. J. Smith, P. Z. T. Haustein, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[3] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021," *White Paper*, March 2017.

[4] 3GPP, "Study on New Radio (NR) Access Technology - Physical Layer Aspects," TR 38.802 (Rel. 14), 2017.

[5] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[6] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, March 2014.

[7] 3GPP, "Study on Integrated Access and Backhaul for NR," AT&T, Qualcomm, Samsung - Tdoc RP-171880, 2017.

[8] H. S. Dhillon and G. Caire, "Wireless backhaul networks: Capacity bound, scalability analysis and design guidelines," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6043–6056, Nov 2015.

[9] M. Polese, M. Giordani, A. Roy, D. Castor, and M. Zorzi, "Distributed Path Selection Strategies for Integrated Access and Backhaul at mmWaves," *IEEE Global Communications Conference (GLOBECOM)*, 2018.

[10] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "How Many Hops Can Self-Backhauled Millimeter Wave Cellular Networks Support?" *arXiv preprint arXiv:1805.01040*, 2018.

[11] M. Mezzavilla, M. Zhang, M. Polese, R. Ford, S. Dutta, S. Rangan, and M. Zorzi, "End-to-End Simulation of 5G mmWave Networks," *IEEE Communications Surveys Tutorials*, Apr 2018.

[12] V. Gambiroza, B. Sadeghi, and E. W. Knightly, "End-to-end performance and fairness in multihop wireless backhaul networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '04. ACM, 2004, pp. 287–301.

[13] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Relay radio transmission and reception," TR 36.826 (Rel. 11), 2012.

[14] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. E. Fuja, "Bandwidth-and power-efficient routing in linear wireless networks," *IEEE Trans. on Inf. Theory*, vol. 52, no. 6, pp. 2624–2633, June 2006.

[15] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 12, pp. 94–101, December 2008.

[16] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2196–2211, May 2015.

[17] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," *IEEE Network*, vol. 28, no. 6, pp. 6–11, Nov 2014.

[18] A. Mesodiakaki, A. Kassler, E. Zola, M. Ferndahl, and T. Cai, "Energy efficient line-of-sight millimeter wave small cell backhaul: 60, 70, 80 or 140 GHz?" in *IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2016.

[19] C. Saha, M. Afshang, and H. S. Dhillon, "Integrated mmWave Access and Backhaul in 5G: Bandwidth Partitioning and Downlink Analysis," in *IEEE International Conference on Communications (ICC)*, May 2018.

[20] 3GPP, "Study on integrated access and backhaul," TR 38.874 (Rel. 15), 2018.

[21] 3GPP, “Service requirements for next generation new services and markets,” TS 22.261 (Rel. 15), 2018.

[22] S. Dutta, M. Mezzavilla, R. Ford, M. Zhang, S. Rangan, and M. Zorzi, “Frame structure design and analysis for millimeter wave cellular systems,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1508–1522, March 2017.

[23] M. Alicherry, R. Bhatia, and L. E. Li, “Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks,” in *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom ’05. ACM, 2005.

[24] 3GPP, “Way Forward - IAB Architecture for L2/3 relaying,” Qualcomm Inc, KDDI, AT&T, Nokia, Nokia Shangai Bell, Huawi, Ericsson, Intel, LG Electronics, CMCC, Samsung - Tdoc R3-181502, 2018.

[25] 3GPP, “NG-RAN; Architecture description,” TS 38.401 (Rel. 15), 2018.

[26] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, “Initial Access Frameworks for 3GPP NR at mmWave Frequencies,” *17th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)*, 2018.

[27] ——, “A Tutorial on Beam Management for 3GPP NR at mmWave Frequencies,” *submitted to IEEE Communications Surveys & Tutorials*, 2018. [Online]. Available: https://arxiv.org/abs/1804.01908

[28] M. Rebato, L. Resteghini, C. Mazzucco, and M. Zorzi, “Study of Realistic Antenna Patterns in 5G mmWave Cellular Scenarios,” in *IEEE International Conference on Communications (ICC)*, May 2018.