# 5G-Flow: A unified Multi-RAT RAN architecture for beyond 5G networks

Meghna Khaturia [a],[*], Pranav Jha [a], Abhay Karandikar [a],[b]

[a] *Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India*
[b] *Director, Indian Institute of Technology Kanpur, India*

## ARTICLE INFO

## ABSTRACT

Convergence of multiple access technologies is one of the key enablers in providing diverse set of services to the Fifth Generation (5G) users. Though 3rd Generation Partnership Project (3GPP) 5G standard defines a common core supporting multiple Radio Access Technologies (RATs), Radio Access Network (RAN) level decisions are taken separately across individual RATs as the existing 5G architecture lacks a unified control and management framework for a multi-RAT network. A unified access network is likely to utilize RAN resources more efficiently and provide an improved performance. In order to achieve a unified Multi-RAT RAN, we supplement the existing 3GPP 5G RAN architecture with OpenFlow. We refer to the proposed architecture as 5G-Flow. The proposed architecture can be viewed as a step towards further evolution of mobile networks beyond 5G. With minimal changes in the 3GPP 5G RAN and none in the core network, we are able to realize a unified and integrated multi-access 5G-Flow RAN. We simplify the existing 5G RAN by replacing RAN nodes with OpenFlow switches and a software-defined controller. The proposed architecture also allows us to completely decouple User Equipment's (UE's) communication with Core Network (CN) from its communication with RAN enabling a UE to use any RAT to connect to any CN (say, use 5G New Radio access to connect to 4G CN, which is not possible in the 3GPP architecture) or to directly connect to Internet from RAN without going via the CN. We have developed an evaluation platform to compare the performance of our proposal with the standard 3GPP architecture. Results demonstrate significant gains in the network performance of 5G-Flow RAN over the existing 3GPP 5G network.

## 1. Introduction

The Fifth Generation (5G) cellular standard envisions to support various use-cases, i.e., enhanced mobile broadband, ultra-reliable low latency communications, and massive machine-type communications [1]. These new requirements have led to several innovations and a redesign of the cellular network as part of 3rd Generation Partnership Project (3GPP) 5G standardization, such as service-based architecture, virtualized network functions, control and user planes separation, and network slicing. One of the major advancements is the unification of multiple access technologies in 5G Core Network (5GC), which is an essential step towards enhancing network efficiency. The 3GPP 5G access network incorporates heterogeneous access technologies, i.e., 3GPP access (e.g., New Radio (NR) based next-generation NodeB (gNB), and Long Term Evolution (LTE) based evolved NodeB (eNB)) and non-3GPP access (e.g., Wi-Fi and wireline access). Towards this, 5G standard has introduced a common interface between 5GC and 5G Radio Access Network (RAN), to integrate multiple access networks at the core network. However, there are certain gaps in the current 3GPP 5G architecture that we discuss next.

Even though 5GC exposes a common interface towards RAN irrespective of the Radio Access Technology (RAT) being used, i.e., N2 interface for control signaling and N3 interface for user data transfer, 5G standard defines different inter-working entities (between RAN and 5GC) for different RATs [1]. For instance, a network entity called Non-3GPP Inter-Working Function (N3IWF) has been introduced for untrusted non-3GPP access (e.g., Wi-Fi), which incorporates N2/N3 protocol stack. In contrast, for 5G NR based access, N2/N3 protocol stack is incorporated within gNB, a new node providing NR data and control planes protocol terminations towards the UE. RAT-specific inter-working entities and associated interfaces result in a complex RAN architecture. Complexity also increases at 5GC, as registration state per RAT is maintained at 5GC if a User Equipment (UE) is connected to more than one RAT. 5GC also controls traffic steering and switching across RATs for UEs concurrently connected to multiple RATs [1,2]. However, 5GC may be unable to optimally manage the traffic across RATs as it has limited RAN level information. Another feature of the 3GPP 5G architecture is the tight coupling between RAN and CN which results in an inflexible architecture. For instance, it is not possible

---

* Corresponding author.
  *E-mail address:* meghnakhaturia@gmail.com (M. Khaturia).

to use 5G NR based gNB with Fourth Generation (4G) Core Network (CN) in the 3GPP 5G network. The current 5G Non-Standalone (NSA) architecture can be viewed as an implication of this feature, wherein 4G eNB has to deployed alongside gNB to utilize the existing 4G CN infrastructure.

Taking into consideration the above-mentioned limitations, we re-architect the 3GPP 5G RAN to realize a unified, integrated, software-defined multi-RAT RAN using OpenFlow Protocol [3]. We refer to the proposed RAN architecture as *5G-Flow*. It employs OpenFlow switches (network switches based on OpenFlow protocol) and a light-weight OpenFlow controller (also called 5G-Flow controller). Unlike 3GPP 5G RAN design, we propose a *unified inter-working interface* towards 5GC. To realize this entity, we split the protocol stack at RAN nodes (i.e., gNB, N3IWF, etc.), separating the radio protocol stack from the N2/N3 protocol stack. We bring together the radio protocol stack of different RATs and N2/N3 protocol stack as different interfaces of an OpenFlow switch in the multi-RAT network. Similarly, OpenFlow switch at UE decouples Non-Access Stratum (NAS) and IP layers from the underlying radio interface stack at the UE. In essence, the proposed protocol split at UE and RAN nodes along with the usage of OpenFlow switches (to connect the protocol stacks of these entities) decouples UE's communication with CN from its communication with RAN. This feature allows a UE to flexibly use any radio interface to connect to any CN. For instance, a UE can use 5G NR based radio interface to connect to 4G CN or directly connect to Internet bypassing the CN, which is not possible in the existing 3GPP 5G network. We summarize some of the novel contributions of the work presented in this paper —

- One of the key innovations is to integrate multiple RATs and realize a *unified inter-working interface* towards 5GC using an OpenFlow switch.
- A light-weight 5G-Flow controller enables RAN level management of downlink as well as uplink dataflows, in turn, utilizing the multi-RAT resources efficiently. Existing OpenFlow abstractions, such as, logical ports and flows are used to control the user traffic flowing through the OpenFlow interfaces.
- 5G-Flow controller manages multiple RATs through OpenFlow protocol messages. The interfaces at OpenFlow switches are responsible for translating these messages to interface-specific action messages.
- UE's connectivity to RAN is fully decoupled from its connectivity to CN. This brings immense flexibility and enables a UE to interface with 5GC, 4G CN, Internet, or any other data network via any 4G/5G/Wi-Fi based RAN.
- Decoupling of radio and CN protocol stacks at RAN allows us to support the 5G non-standalone architecture in a simplified manner wherein a UE need not be dual connected to gNB and eNB. It can connect to 4G CN via gNB.
- We have developed a 5G simulator with Multi-RAT support (for NR and Wi-Fi) since an open source 5G simulator is not yet available. The source code of the 5G multi-RAT simulator has been released under MIT License for the benefit of the research community.
- We have also designed a heuristic algorithms for data flow management for performance evaluation of the proposed 5G Flow network.

As would become clear later, we do not propose to replace the existing protocols and interfaces defined by the 3GPP 5G standards. Instead, the proposal intends to supplement the existing 5G standards to achieve a unified multi-RAT RAN by standardizing those interfaces which are typically deemed proprietary by 3GPP. The proposal requires minimal changes in RAN and none in CN.

The rest of the paper is organized as follows. In Sections 2 and 3, we discuss the literature and background related to the proposed work. In Sections 4 and 5, we discuss the proposed network architecture and its working. We discuss the performance analysis and applications of our proposal in Sections 6 and 7, respectively. We conclude our work in Section 8.

## 2. Related work

Several efforts in the literature deal with the application of SDN in cellular networks to make the network more programmable and in turn, more manageable. Research works such as SoftRAN [4], SoftMobile [5], FlexRAN [6], SoftNet [7] and Software-Defined Wireless Network (SDWN) [8] aim at making RAN more programmable, flexible and specifically, use resources efficiently. SoftAir [9] uses SDN paradigm and hardware abstraction to enable a scalable network architecture for 5G. SoftCell [10] aims to introduce scalability and flexibility in cellular core networks. To support rapid protocol evolution, OpenRadio uses SDN paradigm to decouple the wireless protocol into processing and decision planes and provide a programmable interface between the two planes [11]. However, none of the above works consider multi-RAT integration in the access network.

We now discuss some research works that propose SDN based network architecture to manage heterogeneous RATs. In [12], the authors introduce a virtual MAC layer at network nodes and users in order to manage heterogeneous RATs in a technology independent manner. In [13,14], the authors have developed a prototype that augments WiMAX and Wi-Fi APs with OpenFlow. OpenRAN [15] uses virtualization to design a software-defined multi-access RAN. However, extension of the proposed concept in [12–15] for 3GPP 4G/5G network is not straightforward. 5G-EmPOWER [16] proposes an SDN based framework to manage heterogeneous access networks including Wi-Fi and 4G. However, we do not intend to propose a new framework but suggest enhancements to the existing 3GPP 5G architecture to realize a multi-RAT RAN.

The authors in [17] propose an end-to-end SDN based architecture for a mobile network with a controller managing both access and CN data plane elements by providing an abstract view of the underlying network resources to the controller. In [18], the authors propose a multi-RAT architecture and define a unified and open interface between the controller and multi-RAT data plane. In [19], the authors propose a convergence sub-layer over layer 2 of multiple RATs in order to tightly integrate them. While [17–19] modify the radio protocol stack for RAT integration, the radio stack of different RATs in 5G-Flow RAN remains unchanged. To integrate LTE with Wi-Fi, LTE-Wireless Local Area Network Aggregation (LWA) mechanism has been proposed by 3GPP [20,21]. LWA proposes an emulation layer over Wi-Fi AP that encapsulates LTE packets in Wi-Fi MAC frame.

To the best of our knowledge, none of the above works propose an integrated Multi-RAT RAN with a unified CN inter-working entity. Moreover, the usage of OpenFlow to achieve the same, with minimal modifications in the existing 3GPP protocol stacks and interfaces, has also not been discussed in the literature. Another novelty of our proposal is the decoupling of RAN from the CN and treat UE-CN communication as an overlay communication over the UE-RAN (radio) link, which allows a UE to flexibly use any radio link (NR/LTE/Wi-Fi) to connect to any CN (5G/4G) or even to Internet directly. This also has not been discussed in the literature as yet.

## 3. Background

In this section, we discuss OpenFlow protocol and some basic 3GPP 5G terminologies which form the basis of our discussion ahead.

**OpenFlow Protocol**: OpenFlow is a protocol used by an SDN Controller to manage the forwarding plane of a network. [3]. An OpenFlow compliant network consists of a logically centralized OpenFlow controller (SDN controller) and multiple OpenFlow switches (forwarding plane). An OpenFlow switch supports an OpenFlow client, which communicates with an OpenFlow controller through a secure channel. The controller manages the flow tables present in OpenFlow switches using OpenFlow protocol. A flow-table comprises several flow-entries that match on a packet based on match-fields such as IP address or TCP
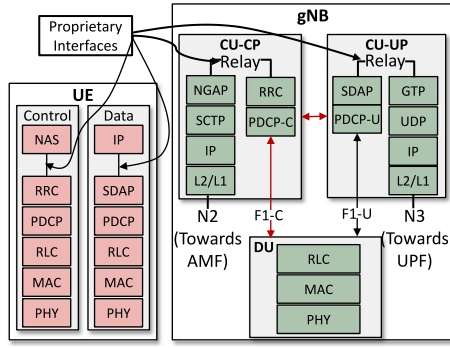
**Fig. 1.** 3GPP 5G gNB and UE protocol stack.

port number. Based on the flow-entry, an action is taken (e.g. forward or drop) on the matched packet.

An OpenFlow switch supports physical and logical ports. A physical port corresponds to a hardware interface on an OpenFlow switch. A logical port does not directly correspond to a hardware interface, but it is an abstraction of a physical port. It can be used to implement processing of packets in OpenFlow switch. *In our proposal, the complex protocol processing is done as part of the logical port in the OpenFlow switch.* OpenFlow-Configuration (OF-Config) protocol, a complementary protocol based on Network Configuration Protocol (NETCONF) [22], helps in configuration and management of OpenFlow switches [23]. OF-Config is responsible for association between a controller and an OpenFlow switch and configuration of physical and logical ports. To inform the OpenFlow controller about various events at OpenFlow switches such as link failures or configuration changes, OpenFlow Notifications Framework provides a method to subscribe to asynchronous messages based on predefined events [24].

**5G Primer**: Fig. 1 illustrates the protocol stack of UE and gNB as defined in the 3GPP 5G standard [1]. gNB consists of a Centralized Unit (CU) and one or many Distributed Units (DUs). gNB-CU can be further divided into control plane (gNB-CU-CP) and user plane (gNB-CU-UP) components. UE and gNB both have Radio Resource Control (RRC) layer that facilitates control plane interaction between the two entities. RRC also helps in radio connection establishment and release, mobility management, and setting up of radio bearers. Service Data Adaptation Protocol (SDAP) layer, along with the underlying protocol stack, at UE and gNB, is responsible for user plane data transfer over the radio interface along with Quality of Service (QoS) handling of dataflows. NAS Layer present at UE is responsible for non-radio related signaling between UE and 5GC. 5G RAN communicates with 5GC through N2 and N3 interfaces. Next-Generation Application Protocol (NGAP) layer, together with the underlying protocol stack (N2 interface), is responsible for all signaling (control) message exchange between RAN and Access and Mobility Management Function (AMF) in 5GC. Data packets between RAN and User Plane Function (UPF) are exchanged over N3 interface using GPRS Tunneling Protocol (GTP) and the underlying UDP/IP protocol stack.

The NGAP-RRC and SDAP-GTP protocol interfaces in gNB are tightly coupled with each other and utilize proprietary vendor specific communication mechanism, as shown in Fig. 1. Though not shown in the figure, other 5G RAN nodes such as N3IWF also use proprietary interfaces for communication between their radio and CN stacks. *The tight coupling between radio and CN protocol stacks on RAN nodes is one of the reasons behind RAT specific Inter-working functions in 5G.* Similarly, higher layers (NAS and IP Layers) at UE are tightly coupled to the underlying radio protocol stack, leading to close inter-working between RAN and CN.

When a UE wants to connect with a data network, such as Internet, the 5G network establishes an end-to-end tunnel between UE and

UPF (i.e., a Protocol Data Unit (PDU) Session). Additionally, a unique signaling link between UE and 5G Network (both at 5GC and RAN) is also established to exchange control messages between the UE and the 5G network. The end-to-end data tunnel (PDU Session) takes the form of a Data Radio Bearer (DRB) at the air interface, while the signaling link takes the form of a Signaling Radio Bearer (SRB). Radio bearers are essentially layer 2 tunnels. Between gNB and 5GC (AMF+SMF), UE specific signaling messages (NAS messages) are exchanged through a (per UE) unique association over N2 interface, whereas a PDU Session takes the form of a GTP tunnel between gNB and 5GC (UPF).

In 3GPP 5G network, different unique identifiers are used to identify UE specific signaling association and data tunnels over different interfaces. For example, to uniquely identify UE over N2 interface, gNB uses RAN UE NGAP ID, whereas AMF uses AMF UE NGAP ID. Similarly, UE specific data tunnels are uniquely identified via GTP Tunnel End-point Indicator (TEID) on N3 interface.

## 4. 5G-Flow RAN architecture

The current 3GPP 5G multi-RAT RAN architecture is illustrated in Fig. 2a. The figure shows how various access technologies interface with 5GC using separate inter-working entities. Untrusted Wi-Fi access uses N3IWF, trusted Wi-Fi access uses Trusted Non3GPP Gateway (TNGF), and wireline access uses Wireline Access Gateway Function (W-AGF) to interface with 5GC. Fig. 2b illustrates our proposed multi-RAT RAN architecture. A multi-RAT 5G-Flow RAN communicates with 5GC through a *unified inter-working interface* instead of separate inter-working functions. Additionally, we aim to enable a flexible interface between RAN and CN, such that any RAN can connect with any cellular CN or Internet directly. The software-defined *5G-Flow controller* acts as a multi-RAT RAN controller that manages the unified inter-working entity and dataflows across multiple RATs in RAN. Since the controller has access to RAN-level information such as traffic load and radio channel conditions, it can efficiently manage the downlink dataflows across RATs. As shown in Fig. 2b, the controller also controls the UE, which enables uplink dataflow management in a multi-RAT RAN. The figure illustrates communication link between UE and controller via gNB–NR (in-band link). However, any RAT can be used for UE-controller communication. To realize the proposed 5G-Flow RAN[1] architecture, we apply the OpenFlow concepts [3]. We envision 5G RAN as an OpenFlow network comprising a 5G-Flow controller (as OpenFlow controller) and OpenFlow switches that are instantiated on the network side and the associated UEs, as shown in Fig. 3a. We discuss the enhancements to the current 3GPP 5G architecture to realize the 5G-Flow network in the remainder of this section.

### 4.1. Proposed multi-RAT network

The existing 3GPP 5G RAN consists of various multi-RAT network nodes, including 3GPP access (e.g., gNB, eNB) and non-3GPP access (e.g., Wi-Fi, N3IWF). To integrate multiple RATs in 5G-Flow RAN and enable a unified inter-working interface, we propose a protocol split between radio interface and N2/N3 protocol stack of RAN nodes. For 3GPP access nodes such as gNB, split happens at the gNB node itself, whereas for non-3GPP access, it is done at the inter-working function such as N3IWF. To illustrate the protocol split, we take an example of gNB. As discussed previously, gNB consists of tightly coupled NR protocol stack (which interfaces with a UE) and N2/N3 protocol stack (which interfaces with the 5GC). As shown in Fig. 3b, we split gNB vertically and separate gNB–NR and N2/N3 protocol stack. We can similarly split N3IWF node. We now introduce an OpenFlow switch, referred to as Multi-RAT Network (MRN) OpenFlow switch, which is

---

[1] Although we discuss radio access technologies in the rest of the paper, the network architecture is applicable to wireline technologies as well.
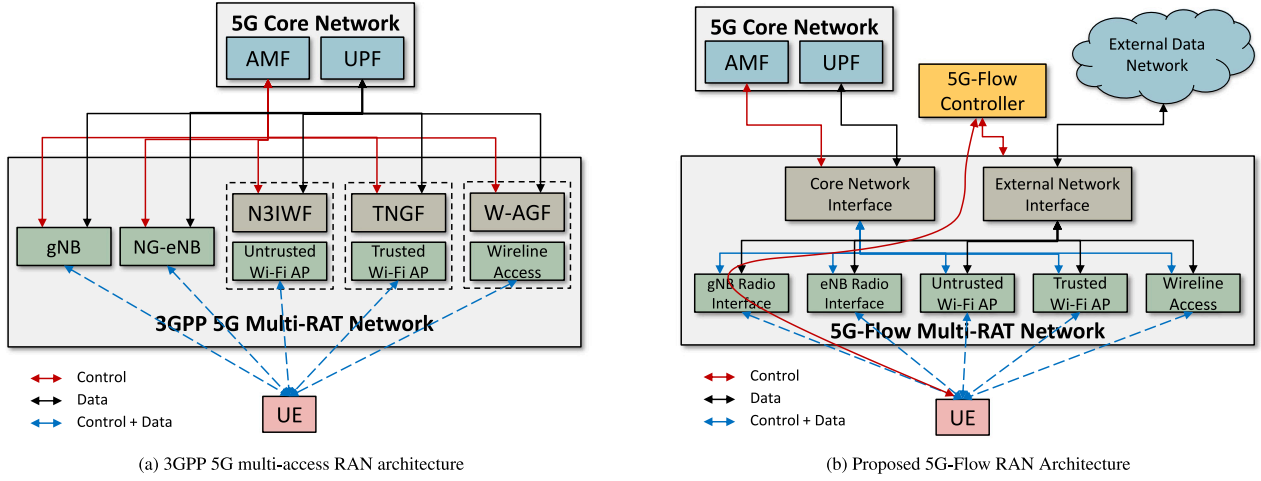
Fig. 2. Conceptual diagram for 5G-Flow RAN architecture and its comparison with the current 3GPP 5G RAN.

(a) 3GPP 5G multi-access RAN architecture

(b) Proposed 5G-Flow RAN Architecture



(a) Block Diagram of 5G-Flow RAN

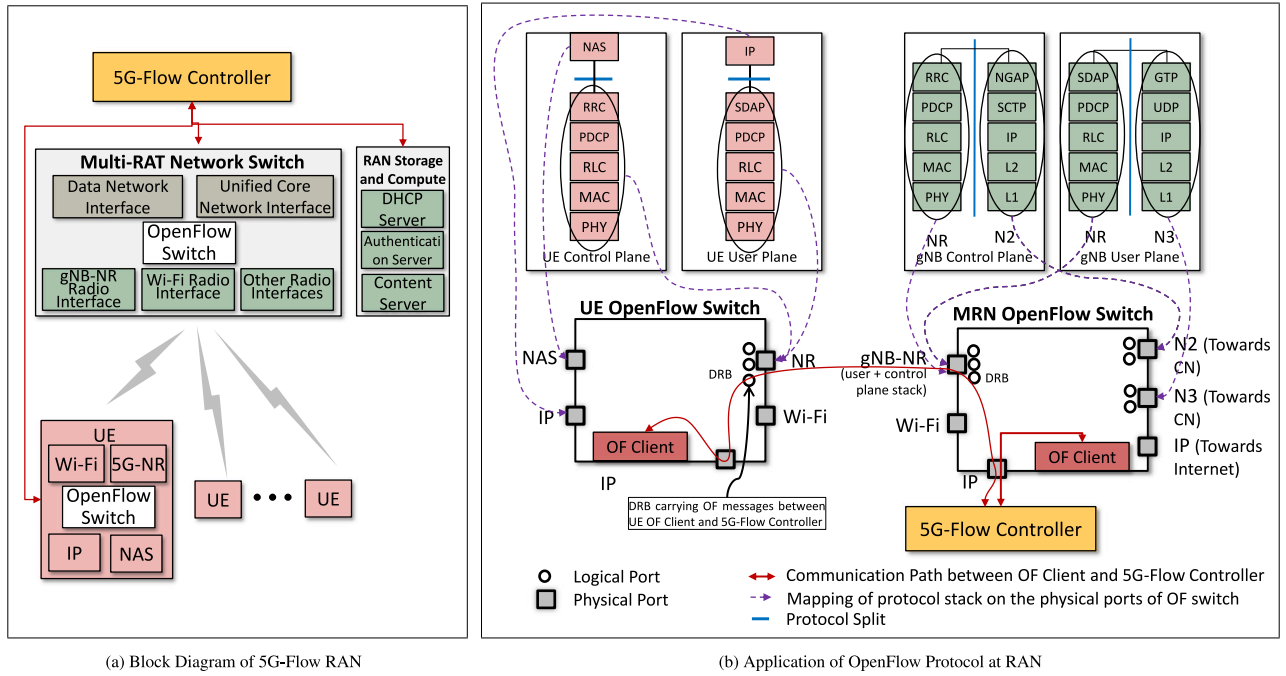(b) Application of OpenFlow Protocol at RAN

Fig. 3. Implementation of 5G-Flow RAN architecture using OpenFlow Protocol.

responsible for bridging radio and N2/N3 protocol stacks of multiple RATs. These protocol stacks form different interfaces (physical ports) of MRN OpenFlow switch. The 5G-Flow controller directs MRN OpenFlow switch to process the messages from different radio interfaces and deliver them to N2/N3 stack towards 5GC and vice-versa. This way, the MRN OpenFlow switch along with the 5G-Flow controller replaces all RAT-specific entities such as gNB, eNB, N3IWF, etc. and exposes a unified interface towards CN.

The MRN OpenFlow switch has physical ports both at the radio and the 5GC interface side, as shown in Fig. 3b. Both the control plane (RRC and underlying protocol stack) and the data plane (SDAP and underlying protocol stack) of gNB–NR radio interface map to one of the radio side ports of the OpenFlow switch. Similarly, Wi-Fi Media Access Control (MAC) and physical layer map to another port. NGAP and GTP protocol layers (along with the underlying N2/N3 stack) map to the physical ports on the 5GC side. The physical port, labeled as IP, interfaces with the external data network. As the radio interface stack is decoupled from the N2/N3 protocol stack, OpenFlow switch can steer the data traffic of a UE towards IP port, enabling direct connectivity with Internet bypassing the CN. An interface towards 4G CN is not shown in the figure, but it can be easily incorporated by adding S1-Mobility Management Entity (S1-MME) interface as a separate physical port in the proposed OpenFlow switch. This feature enables a UE with 4G compatible NAS layer to communicate with 4G CN via 5G RAN. This feature is discussed in the detail in Section 8.1.

In Fig. 3b, we have mapped Wi-Fi MAC layer to the radio side port in MRN OpenFlow switch. However, the user plane of N3IWF uses additional protocol layers i.e., Generic Routing Encapsulation (GRE) and IP Security (IPsec), for creating a secure tunnel between UE and N3IWF over Wi-Fi radio interface. Similarly, it uses TCP and IPsec protocol layers in the control plane for encapsulating NAS messages. Our architecture provides flexibility in employing these protocols. If a UE does not want to use GRE and the underlying protocol stack for some dataflows, a logical port can be created at UE and RAN, which transparently passes the data packets through Wi-Fi interface without any processing. Further, a logical port enabled with GRE, and IPsec

protocol layers can be created over Wi-Fi MAC based physical port if a UE needs a secure tunnel for another data flow.

MRN OpenFlow switch is located at the edge (between the RAN and the core). A single MRN OpenFlow switch can handle multiple RATs such as gNB–NR and Wi-Fi AP. The gNB–NR and Wi-Fi AP RATs may be connected to MRN OpenFlow switch via mid-haul connectivity, and they need not be collocated. To manage a large network, multiple MRN OpenFlow switches can be deployed. The 5G-Flow controller is located at the edge. There can be one or more controllers to handle multiple MRN OpenFlow switches. We plan to study the requirement of an interface between 5G-Flow controllers in the future.

### 4.2. Proposed enhancements at UE

An OpenFlow switch, introduced at UE, decouples NAS (that communicates with 5GC) and IP protocol layers from the underlying radio protocol stack, as shown in Fig. 3b. We introduce a common IP layer instead of RAT-specific IP layers. There can be different radio interface stacks depending on the technology, but NAS and IP layers remain common. At the radio side ports of UE OpenFlow switch, NR (RRC+SDAP, and the underlying protocol stack) and Wi-Fi (MAC and physical layer) radio stacks are mapped.

The UE OpenFlow switch, along with the 5G-Flow controller, manages the UE's radio connectivity and enables uplink dataflow management across multiple RATs. Moreover, when a UE is connected to 5GC via multiple RATs, it registers only once. In the existing 3GPP architecture, a UE connected to 5GC through more than one RAT has to register with 5GC separately via each RAT. With the separation of NAS layer from radio protocol stack (in addition to radio and N2/N3 protocol split at the network side), UE's communication with RAN is completely decoupled from its communication with 5GC. This feature allows a UE to flexibly connect to different networks such as 5GC, 4G CN, or directly to Internet.

### 4.3. OpenFlow switch configuration

As discussed before, OpenFlow controller can create and configure logical ports on an OpenFlow switch using OF-Config protocol. The meaning (processing) associated with these logical ports varies according to the underlying interface being used. On the gNB–NR interface at the MRN OpenFlow switch, a logical port represents a UE-specific radio bearer. At least two logical ports are created for each UE, one for SRB and another for DRB. To identify these logical ports uniquely, we use SRB/DRB ID as defined in 5G standard. A logical port at NGAP interface signifies a UE-specific NGAP association and is identified by RAN UE NGAP ID. Similarly, at GTP interface, a logical port implies a PDU session of a UE and is identified by a GTP-TEID.

At NR interface of the UE OpenFlow switch, a logical port represents SRB and DRB of a UE (similar to MRN OpenFlow switch). At IP interface, logical ports signify ongoing PDU sessions for a UE.

*Despite the differences in processing at each of the physical interfaces, usage of logical ports provides a uniform abstraction to be used by 5G-Flow controller to configure flow paths through the switch.* The 5G-Flow controller simply configures logical ports on the physical interfaces in an OpenFlow switch. It is the responsibility of the interface to translate OF-Config message (for port creation) to interface specific handling. For instance, a DRB, along with a GTP tunnel, needs to be established during a PDU Session setup. When the gNB–NR interface receives a message to create a new logical port corresponding to a DRB, the RRC layer on gNB–NR interface translates it to configure its underlying lower layers, e.g., PDCP, RLC layers for the local DRB configuration. It also exchanges *RRC Reconfiguration* messages with the UE for a corresponding DRB configuration on the UE. Similarly, a logical port creation message sent to N3 interface is translated (by the interface) to create a GTP tunnel. A logical port creation message to N2 interface gets
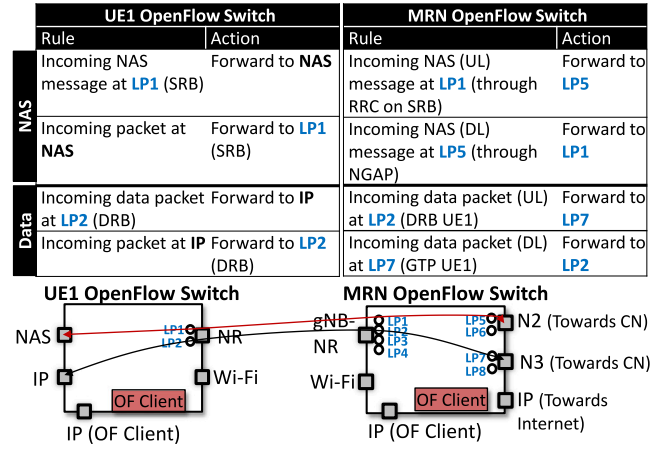


| UE1 OpenFlow Switch | | MRN OpenFlow Switch | |
|---|---|---|---|
| **Rule** | **Action** | **Rule** | **Action** |
| **NAS** Incoming NAS message at LP1 (SRB) | Forward to NAS | Incoming NAS (UL) message at LP1 (through RRC on SRB) | Forward to LP5 |
| Incoming packet at NAS | Forward to LP1 (SRB) | Incoming NAS (DL) message at LP5 (through NGAP) | Forward to LP1 |
| **Data** Incoming data packet at LP2 (DRB) | Forward to IP | Incoming data packet (UL) at LP2 (DRB UE1) | Forward to LP7 |
| Incoming packet at IP | Forward to LP2 (DRB) | Incoming data packet (DL) at LP7 (GTP UE1) | Forward to LP2 |

**Fig. 4.** Example of flow entries at UE and MRN OpenFlow switch for NAS signaling and data transfer. Figure shows how Logical Ports (LPs) are used to set up a flow in the proposed network.

translated into the creation of a unique UE-specific NGAP association with 5GC.

Once the logical ports are created on the interfaces, 5G-Flow controller defines a mapping between different logical ports across the interfaces on the OpenFlow switch, as shown in Fig. 4. For instance, a mapping between DRB and GTP tunnel is created for a UE, or a mapping between UE-specific NGAP association and SRB is created. The controller realizes these mappings through flow entries added at the OpenFlow switches. These port mappings enable the simple forwarding of data and signaling messages through the OpenFlow switch and also makes the control and management task for the 5G-Flow controller easy.

A key novelty of the proposal is the usage of GTP tunnels, radio bearers, or UE-specific NGAP association as logical ports. These entities (a radio bearer, or a UE-specific NGAP association) carry a specific set of data flows or messages. By using them as logical ports, we can virtualize them and enable their manipulation by an OpenFlow Controller through OpenFlow protocol.

## 5. Working of 5G-Flow RAN

### 5.1. Initial connection setup

In this section, we discuss how OpenFlow switches can communicate with the 5G-Flow controller. The control messages between an OpenFlow switch and the controller are exchanged over a TCP/IP connection. The MRN OpenFlow switch and the 5G-Flow controller are co-located and can communicate over a wired interface. Fig. 5 shows how communication between an UE OpenFlow Client (UE OpenFlow switch) and the 5G-Flow controller is established. Here, we have assumed that the default path for UE-controller communication is via gNB–NR interface, but the path can be established via Wi-Fi radio interface as well. We explain the call flow to set up the initial connection next.

- To establish a radio connection with gNB–NR interface, UE sends an *RRC Setup Request* over common channel (SRB0). Since 5G-Flow controller is responsible for admission control, gNB–NR interface notifies the controller via a NETCONF notification so that a decision for UE admission can be taken. If admitted, the 5G-Flow controller sends an OF-Config message to create a logical port on gNB–NR interface for the subsequent signaling messages exchange (via SRB1) with the UE. gNB–NR interface sets up SRB1, maps it to the logical port, and sends *RRC Setup*. As the UE
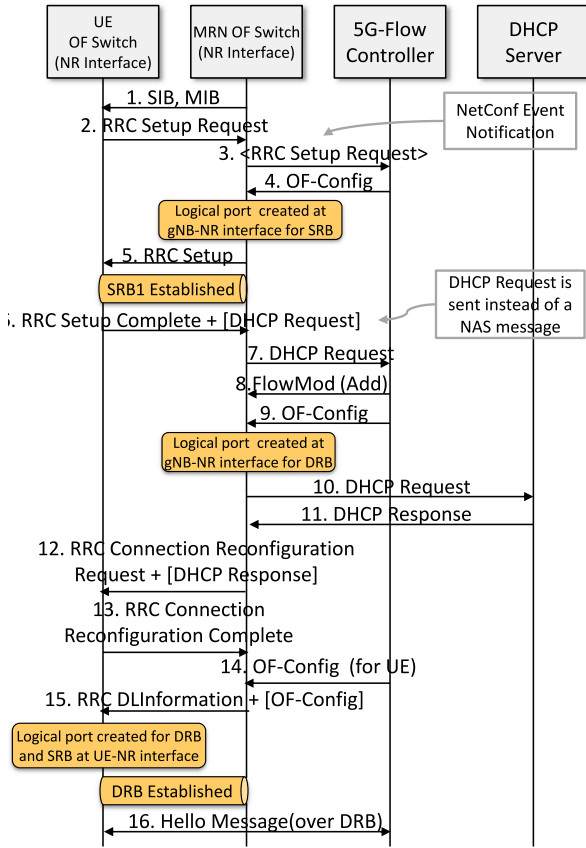
**Fig. 5.** Initial connection setup.

OpenFlow client still does not have a TCP/IP connection with the controller, a logical port cannot be created at UE using OF-Config. Instead, UE uses the default physical port mapped to 5G-NR interface for initial signaling.

- UE responds with *RRC Setup Complete* message and a *DHCP request* is sent in NAS message field instead of *Registration Request*. gNB–NR interface forwards *DHCP request* towards the 5G-Flow controller as a table-miss is triggered at MRN OpenFlow switch. The controller upon receiving this message sends a FlowMod (Add) command to add the flow entry at MRN OpenFlow switch. 5G-Flow controller also sets up a logical port at gNB–NR interface for DRB of the UE, using OF-Config message. This DRB is created for carrying OpenFlow client messages from UE to the 5G-Flow controller.

- gNB–NR interface forwards *DHCP request* to the DHCP server which processes the request. The DHCP response is sent to UE in the NAS message field of *RRC Connection Reconfiguration Request*. A DRB has now been established.

- An OF-Config message is sent to the UE to configure a logical port and map it to the newly established DRB. This message creates a logical port for SRB as well, which is used for future signaling messages. The communication path between the UE OpenFlow switch and the 5G-Flow controller is now established via DRB, and they exchange *Hello* message over this DRB.

This callflow illustrates how UE's communication with RAN is decoupled from its communication with CN. UE can use RAN to exchange DHCP and OpenFlow messages with entities located in the edge instead of exchanging NAS messages and data with the CN.

## 5.2. 5G registration and session setup

If a UE wants to access the 5G cellular network, it needs to register with the 5GC. Here, we take an example of how a UE registers via gNB. The registration of a UE can also take place via a Wi-Fi network, if available. To facilitate UE's communication with the 5GC, 5G-Flow controller can proactively set up a path that delivers NAS messages from a UE to the 5GC and vice-versa, as shown in steps 1 and 2 in Fig. 6. It can also be implemented through a reactive method in which the 5G-Flow controller adds the flow entries after the first NAS message originates at UE, and a table-miss is detected.

Next, we discuss how a PDU session is established via 5G-Flow RAN. When AMF (in 5GC) receives the NAS message (*PDU Session Establishment Request*) from a UE, it informs SMF, which creates a PDU session for the UE. SMF sends PDU session related information to AMF, i.e., GTP tunnel end-point and QoS information. AMF forwards this information as an NGAP message to RAN along with a NAS message (*PDU session Establishment Accept*). The MRN OpenFlow switch transparently forwards the NAS message to the UE. However, unlike NAS message, it processes the NGAP message and forwards PDU session related attributes to the 5G-Flow controller via *OpenFlow Packet-in* message. For processing the NGAP message and enabling communication between gNB–NR interface stack and 5G-Flow controller, a new protocol definition may be required. However, the protocol will be required for limited tasks such as processing and exchanging PDU Session (data bearer) related parameters received/sent via NGAP messages, while most tasks are handled via logical port processing in OpenFlow switch. The controller, based on the Packet-in message and RAT-specific information, decides how the incoming dataflow should be distributed among the available RATs. It, then, sends OF-Config messages to the radio interfaces to create logical ports, which are interpreted by these interfaces for configuration of the underlying protocol stack. For instance, configuration of lower layers by RRC to create a DRB when a logical port creation message is received at gNB–NR interface. Also, an OF-config message is sent to GTP interface of the MRN OpenFlow switch for creating a logical port which signifies GTP tunnel for a UE's PDU Session. After this, flow entries are added at the UE OpenFlow switch and the MRN OpenFlow switch that maps the newly created DRB to IP port at UE and maps the DRB to the GTP tunnel, respectively. The path for UE PDU session is now set up.

## 5.3. Dynamic dataflow management

5G has introduced ATSSS feature, which manages the multi-access traffic in the 5G network. Fig. 7a illustrates how the multi-access traffic flows through the 3GPP 5G network [1,2]. UE can initiate a Multi-Access-Protocol Data Unit (MA-PDU) session to enable PDU exchange between UE and UPF via 3GPP and non-3GPP RATs simultaneously. To manage the uplink traffic, UE considers the ATSSS rules provided by the 5GC. To manage the downlink traffic, UPF considers ATSSS rules provided by SMF over N4 interface along with the feedback information from UE. However, the feedback information available at UPF is limited to only Round Trip Time (RTT) and RAT availability. This information may not be sufficient to route the flow through multiple RATs optimally.

As shown in Fig. 7b, the flow configuration in the proposed 5G-Flow network happens at the RAN-level, where utilizing various RAT-specific attributes is viable. The controller can periodically access the value of specific attributes, such as traffic load at different RAT interfaces, flow statistics, and radio signal strength, to optimize the dataflow. The first two attributes are available at the OpenFlow switch. To access the radio signal measurement available at RRC (gNB–NR), the controller can subscribe to the measurement data at the OpenFlow switch via OF-Config (NETCONF) notifications [24]. The asynchronous notification mechanism is supported by OF-Config protocol and it allows us to
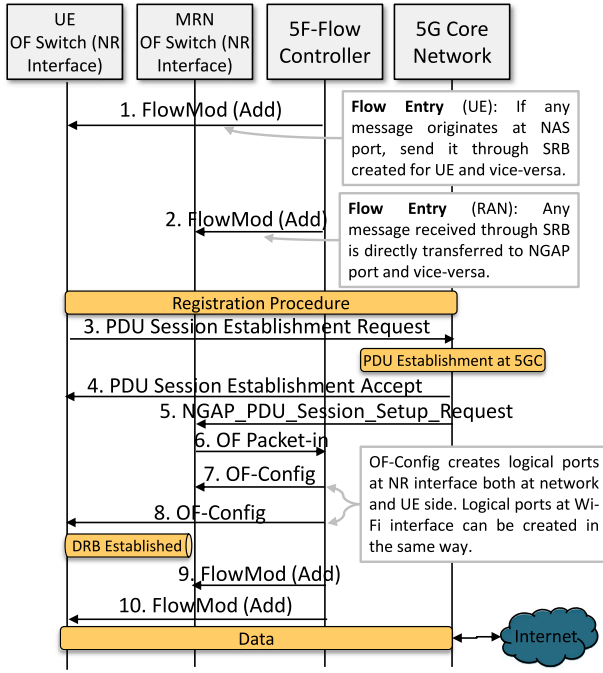
Fig. 6. Call flow to illustrate the communication of UE with 5G Core Network.



(a) Dataflow in 5G Network under ATSSS feature
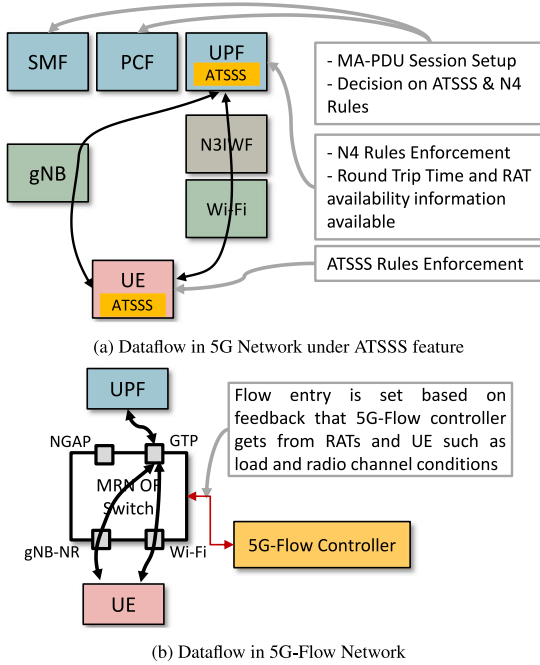


(b) Dataflow in 5G-Flow Network

Fig. 7. Comparison of dataflow with 5G ATSSS feature and dataflow with 5G-Flow architecture.

define notification messages, in addition to the already available set of notifications.

Based on these parameters, an optimized policy for flow management can be determined, and the incoming flow is split across available RATs, as shown in the figure. We also have the flexibility to decouple uplink and downlink in 5G-Flow RAN. As the controller manages MRN and UE OpenFlow switch separately and can add distinct flow entries in both the switches, uplink and downlink for a user can be easily decoupled. We evaluate the performance of data flow management in the 5G-Flow network viz-a-viz the 3GPP 5G network in the next section.
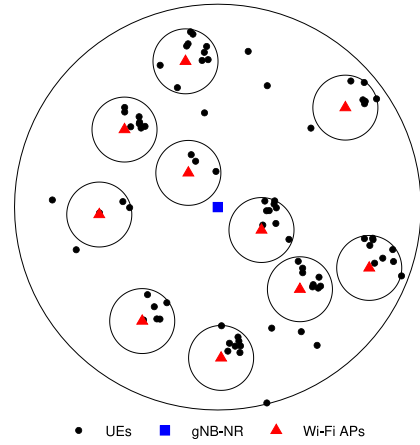


Fig. 8. Single cell simulation scenario.

## 6. Performance analysis

The proposed 5G-Flow architecture enables a framework wherein a multi-RAT RAN can be efficiently managed by the 5G-Flow controller. In this paper, we focus on analyzing the performance of data traffic management across multiple RATs in a 5G-Flow network. To analyze the performance of our architecture, we have built an evaluation platform. An open-source 5G simulator with a fully developed protocol stack is not yet available, so we have developed a 5G simulator and added support for multiple RATs. We have released the source code of the 5G multi-RAT simulator under MIT License for the benefit of other researchers and it is available online [25]. The results presented in this paper can be verified by the researchers and further enhancements to the simulator, if any, can be suggested. In our simulator, packets are the fundamental objects, and we have implemented physical and MAC layer protocol stacks for 5G-NR and Wi-Fi RATs. Support for higher layer protocol stacks such as RRC and NGAP is not added at present. However, we have implemented a centralized controller that manages these RATs.

We perform Monte Carlo simulations for 1 s and average the results over 50 deployment scenarios. We analyze the performance of average TCP throughput and average packet delay for the 5G-Flow network and compare them with the current 3GPP 5G network. In our simulation model, the TCP throughput is the sum of TCP throughput of all the users in a cell. We define packet delay as the total time it takes for a packet to reach its destination node from the source node. The packet delay essentially indicates queuing delay at the access network and transmission delay.

### 6.1. Single cell scenario

We consider a single-cell scenario with 250 m radius as shown in Fig. 8. The network model comprises a 5G NR cell and multiple Wi-Fi Basic Service Areas (BSAs). The cell consists of a gNB–NR entity located at the center. The 5G-Flow controller and the MRN OpenFlow switch are located at the edge near the gNB–NR entity. Multiple Wi-Fi APs provide radio coverage (BSA) inside gNB–NR cell and are distributed uniformly at random. A UE in the network is assumed to have two radio interfaces: 5G-NR and Wi-Fi. We also assume that 80% of users are connected to both the RATs while the remaining users are connected to only gNB–NR entity. The users, in our simulation model, are assumed to be stationary.

For 5G-NR, we use 3GPP Urban Macro (UMa) path loss model, whereas Urban Micro (UMi) path loss model for Wi-Fi [26,27]. We also consider log-normal shadow fading for 5G-NR and Wi-Fi RATs and their standard deviations are as per the UMa and UMi path loss model. The Wi-Fi network is based on IEEE 802.11n [28]. Other important parameters used in our simulator are given in Table 1.

**Table 1**
Network model.

| Parameters | Values |
|---|---|
| Number of Wi-Fi APs | 10 |
| Number of users | 80 |
| Packet payload | 1000 bytes |
| TCP header | 60 bytes |
| 5G-NR network model | |
| Carrier frequency | 1.9 GHz (TDD) |
| 5G numerology | 1 |
| Bandwidth | 60 MHz (162 PRBs, 2 slots per sub-frame) |
| UE/gNB–NR transmit power | 43/23 dBm |
| UE/gNB–NR antenna gain | 2/15 dBi |
| UE/gNB–NR antenna height | 1.5/25 m |
| UE/gNB–NR noise figure | 7/10 dB |
| Wi-Fi network model | |
| Bandwidth | 20 MHz |
| Operating frequency | 2.4 GHz |
| Coverage radius | 40 m |
| UE/AP transmit power | 20/15 dBm |
| AP antenna gain | 4 dBi |
| AP antenna height | 10 m |
| MPDU | 1500 bytes |

**Table 2**
Simulation cases for 80 users in a cell.

| Cases | Service type | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Case a | 0 | 0 | 40 | 40 |
| Case b | 10 | 10 | 30 | 30 |
| Case c | 20 | 20 | 20 | 20 |
| Case d | 30 | 30 | 10 | 10 |
| Case e | 40 | 40 | 0 | 0 |

$l_g$ = Load at gNB

$l_w$ = Load at Wi-Fi AP to which the user is connected

$ch_{g,i}$ = Channel condition for user $i$ under 5G-NR network

$s_i$ = Service type of user $i$

Let $L = \{1, 2, 3\}$ denote the set of values representing low, medium and high load respectively at a particular RAT. The load at gNB–NR ($l_g$) and Wi-Fi AP($l_w$) take its values from $L$. The channel condition experienced by user $i$ from gNB is represented by $ch_{g,i} \in \{0, 1\}$, where 0 represents good channel and 1 represents bad channel. We distinguish between good and bad channel based on a threshold value of received Signal-to-Noise Ratio (SNR), which has been fixed at 6 dB in the simulation model. The service type of user $i$ is represented by $s_i \in \{1, 2, 3, 4\}$, where 1 represents the highest priority service while 4 represents the service type with the least priority. The coefficients in Eq. (1) ($w_k$) are the weights assigned to the metrics in the equation. The weights can be modulated based on the impact of a certain metric on system performance. The weights considered in our algorithm are given in Algorithm 1. We assign highest importance to $ch_{g,i}$ because bad channel leads to resource wastage and poor performance. If the gNB–NR RAT is observing high load, the high value of $w_1$ ensures that additional users are not scheduled. This ensures that the QoS of already scheduled high priority users under gNB–NR is not impaired. After the evaluation of $T_i$, RAT selected for a user $i$ ($R_i$) is determined based on the threshold value $T'$. Since, we aim to balance the load on the available RATs, $T'$ is set to be the mean of all possible values of $T_i$.

$$R_i = \begin{cases} 1, & T_i > T' \quad \text{(1 represents Wi-Fi)} \\ 0, & T_i \leq T' \quad \text{(0 represents gNB)} \end{cases}$$

Since the 5GC is unaware of the RAN level information, the RAT selection policy for standard 3GPP 5G network only considers the service priority. For the performance evaluation of 5G network, we consider that a user with service priority 1 or 2, is always scheduled at gNB–NR and with service priority 3 or 4 is always scheduled at Wi-Fi (irrespective of load or channel condition).

*Results*: Fig. 9 shows the average TCP throughput and packet delay for various simulation cases considered in Table 2. We first analyze the graphs in Figs. 9a and 9e, for extremely skewed service type distribution among users. Since the 5G-Flow network considers load on RATs and channel condition, RAT selection is done in an efficient way. Hence, the throughput and delay performance for 5G-Flow network are significantly better. Since all the users in standard 5G network are scheduled either on Wi-Fi RAT (for Case a in Fig. 9a) or on gNB–NR RAT (for Case e in Fig. 9e), the respective RAT is overloaded and hence the overall performance is poor. This also leads to poor utilization of the other RAT.

Figs. 9b and 9d illustrate average throughput and packet delay results for service type distributions which are relatively less skewed than the cases discussed before. In Fig. 9b, the average packet delay under standard 5G network is quite high for users with service priority 3 and 4 as all the 60 users (with these service priorities) are scheduled at Wi-Fi. gNB–NR under this case, experiences low load and hence the average packet delay is quite low for users with service priority 1 and 2. Similarly, in Fig. 9d, the delay performance of users with service priority 2 suffers as the gNB–NR experiences traffic load from 60 users

## 6.1.1. Downlink dataflow management for UEs with different service types

As discussed before, data flow management involves selecting an optimal RAT (from the available RATs) for each session of a UE based on the RAN-level information. To analyze the performance of 5G-Flow RAN with respect to RAT selection, we consider four types of data services offered in the network in the order of the priority, (i) mission-critical streaming, (ii) real-time streaming, (iii) non-real-time streaming and (iv) best-effort. For simplicity, we assume that a UE requests only one type of service for the entire duration of simulation. We assume Poisson traffic with the packet arrival rate of 500 packets/s. The packet size is fixed at 1000 bytes (payload). Therefore, the downlink bit-rate is 4 Mbps for all the users (irrespective of the service requested), and TCP protocol is used as transport protocol. Since we aim to analyze the downlink performance, we assume that the entire available bandwidth is used by the downlink traffic and do not consider the bandwidth used by TCP-ACK packets (in uplink). We evaluate the uplink performance of the proposed network in the scenario discussed next. We assume 80 users in our simulation model. We consider five different cases based on the service type requested by a user and these are explained in Table 2. The 5G-NR radio resource scheduler uses a priority scheduling algorithm to give better service to higher priority users. Wi-Fi network uses CSMA/CA to schedule the users and does not distinguish between user service priorities.

For the purpose of performance evaluation, we have designed a threshold-based heuristic algorithm for RAT selection given in Algorithm 1. The data related to the metrics or the RAT-specific attributes considered in the algorithm is provided via OpenFlow switches in the 5G-Flow network. As discussed before, the controller can subscribe to measurement data at the OpenFlow switch via NETCONF notifications. The algorithm considers the following metrics. The load conditions at RATs (Wi-Fi and gNB–NR) and 5G-NR radio channel quality are the most critical metrics to select a RAT for UE efficiently. In addition, the RAT selection algorithm also takes into account the type (priority) of service being used. Due to the small coverage area of Wi-Fi AP, received signal strength for UEs do not vary much. Hence, the channel condition for a UE under Wi-Fi network is always considered good. We assign different weights for the different metrics and calculate the value of $T_i$ (for every UE) based on the following equation. We then select the RAT for a UE based on the predetermined threshold value.

$$T_i(l_g, l_w, ch_{g,i}, s_i) = \frac{w_1 \cdot l_g + w_2 \cdot l_w + w_3 \cdot ch_g + w_4 \cdot s_i}{\sum_{k=1}^{4} w_k} \quad (1)$$

where,

**Algorithm 1:** Downlink RAT Selection Algorithm

**Input:**

$C_g = C_0$               `// gNB-NR channel capacity`

$C_w = W_0$            `// Wi-Fi AP channel capacity`

$l_g = 1, l_w = 1$          `// Starting at low load`

$w_1 = 0.26, w_2 = 0.09, w_3 = 0.43, w_4 = 0.22$

$T' = 1.46$

`// N = Number of UEs associated with both RATs`

**for** $i$ *from* 1 *to* $N$ **do**

     Evaluate $ch_{g,i}$ for user $i$

     Calculate $T_i(l_g, l_w, ch_{g,i}, s_i)$ based on Eq. (1)

     **if** $T_i > T'$ **then**

         $R_i = 1$

         Decrease $C_w$

         **if** $C_w = medium$ || $C_w = low$ **then**

             Update $l_w$

         **end**

     **else**

         $R_i = 0$

         Decrease $C_g$

         **if** $C_g = medium$ || $C_g = low$ **then**

             Update $l_g$

         **end**

     **end**

**end**

**Output:** $R_i, \forall i \in \{1, 2, \ldots, N\}$

and it prioritizes the users with priority 1. The performance of evenly distributed user service priority is shown in Fig. 9c. The performances of 5G-Flow network and standard 5G network are comparable as the 5G network follows a RAT selection policy which is best suited for evenly distributed user service priority.
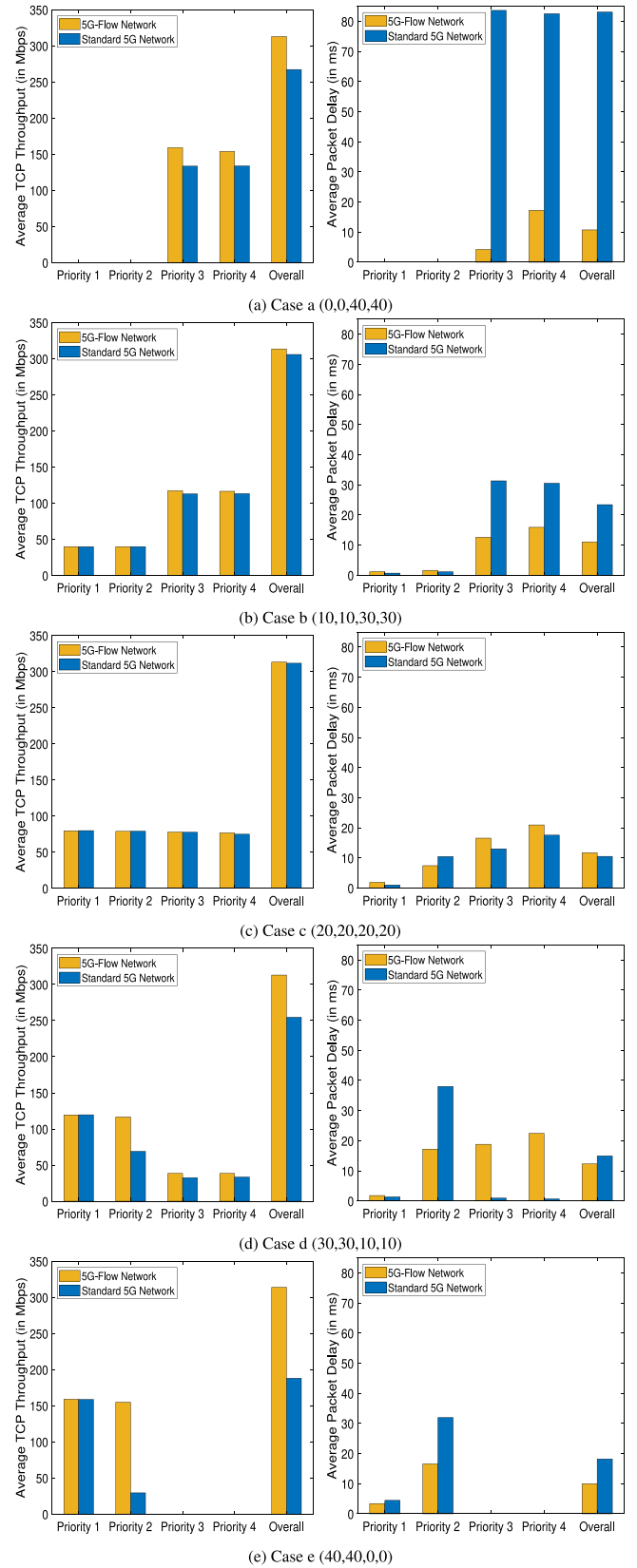
In general, the standard 5G network is unable to efficiently use the available radio resources in a multi-RAT scenario. The performance of the 5G-Flow network significantly improves as it considers RAN-level information while performing RAT selection.

*Load Balancing*: In Fig. 10, the graph shows how the traffic is split across various RATs for simulation cases presented in Table 2. The 5G network in unable to balance the traffic across RATs and hence the overall throughput suffers. On the other hand, 5G-Flow network observes a consistent performance as it balances the load across RATs.

*6.1.2. Uplink dataflow management*

The proposed 5G-Flow network enables flexibility in managing uplink and downlink traffic of every UE independently. Moreover, our architecture decouples uplink and downlink traffic management, i.e., RAT selection for a UE's uplink and downlink traffic is determined independently. Due to this decoupling, we expect that the performance of the overall network and, in particular, of Wi-Fi network will improve when the uplink users are fewer in a Wi-Fi network. This is expected as fewer users will contend to gain access to the channel in a CSMA/CA system.

In the existing 3GPP 5G network, ATSSS feature allows for independent decisions on uplink and downlink traffic distribution. UE selects RAT for uplink transmission based on channel quality that it is observing on multiple RATs and ATSSS rules. UE does not know the load on each RAT to decide the best possible RAT for uplink. For instance, let us assume that a UE observes a good channel gain from Wi-Fi AP as well as gNB–NR, so it selects gNB–NR interface for uplink. gNB–NR may be heavily loaded, and Wi-Fi may be lightly loaded, leading to poor uplink performance (which could have been better on Wi-Fi). Therefore, without loss of generality, we assume that UEs select



(a) Case a (0,0,40,40)

(b) Case b (10,10,30,30)

(c) Case c (20,20,20,20)

(d) Case d (30,30,10,10)

(e) Case e (40,40,0,0)

**Fig. 9.** Downlink performance of 5G-Flow network vis-a-vis 3GPP 5G network.

the same RAT for uplink and downlink i.e., uplink and downlink are essentially coupled for 5G network.
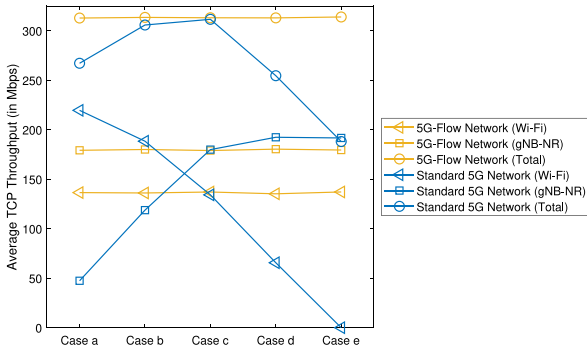
Fig. 10. Load balancing across gNB–NR and Wi-Fi APs.

---

**Algorithm 2:** RAT Selection Algorithm for Uplink

**Input**:

$C_g = C_0$                                  // gNB--NR Uplink channel capacity
/* $W_0$ = Maximum number of users allowed to connect with a
    Wi-Fi AP in uplink                                                     */
$A$ = List of sorted Wi-Fi APs with respect to number of connected UEs

**for** *each i in A* **do**

   $N_i$ = Number of users connected to Wi-Fi AP $i$
   $K = \min(N_i - W_0, C_g)$
   Move $K$ users (closest to gNB) from Wi-Fi to gNB–NR for uplink data transmission
   Update $C_g$

**end**

---
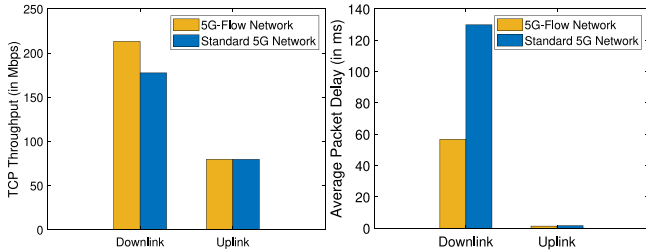


Fig. 11. Performance comparison between 5G-Flow network and Standard 5G network when Uplink and Downlink data flows are decoupled.

In our simulation scenario, we use TDD duplex scheme to allocate resources for uplink and downlink. The transmission periodicity for the TDD radio frame is 5 ms, i.e., the frame repeats after 10 slots (for 30 kHz sub-carrier spacing). TDD frame configuration is {D,D,D,S,U,U,U,U,U,U}, where D, U, and S represent downlink, uplink and special slot respectively [29]. The S slot consists of downlink symbols except for the last symbol, which is used as a switching symbol. We consider 80 users each having 3 Mbps downlink and 1 Mbps uplink data rate requirement. We consider the Poisson traffic model. We do not consider service priority in this simulation scenario. Therefore, we use a round-robin scheduler for 5G-NR to allocate radio resources to the users.

For those users that are dual-connected in the 5G-Flow network, downlink data is scheduled on Wi-Fi. However, we schedule the uplink data to gNB–NR interface for users who are experiencing good channel gain from gNB–NR RAT. The RAT selection algorithm for uplink users is given in Algorithm 2.

*Results*: Here, we consider average TCP throughput for all the users in downlink direction and uplink direction separately. Fig. 11 demonstrates that downlink TCP throughput is higher for the 5G-Flow network as compared to the standard 5G network. We also observe the
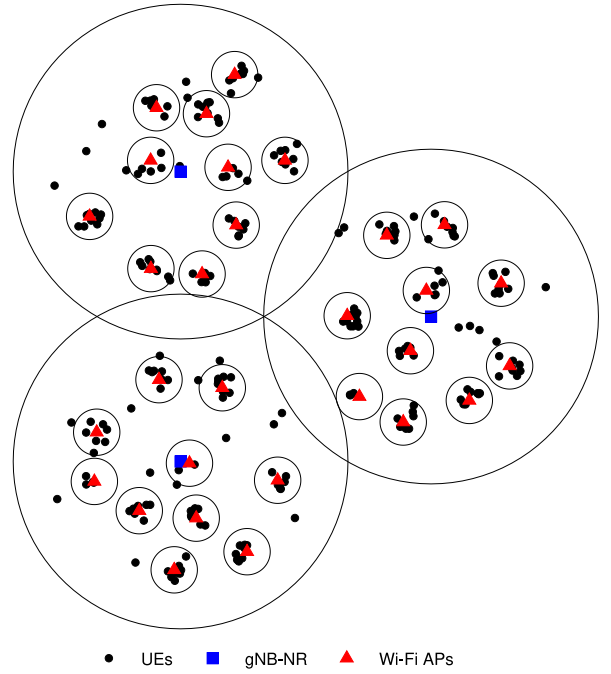


UEs     gNB-NR     Wi-Fi APs

Fig. 12. Multi-cell simulation scenario (Inter-site distance for gNB–NR cells is 500 m).

percentage of served traffic to offered traffic for Wi-Fi APs in our simulations. While it is 75.87% under standard 3GPP 5G network, it is 88.5% for 5G-Flow network. This improvement in 5G-Flow network is observed due to the reduction in load at Wi-Fi APs as we have efficiently managed uplink users based on Algorithm 2. Therefore, fewer users contend to gain access to channel under Wi-Fi APs in 5G-Flow network, thereby increasing performance efficiency of Wi-Fi APs. For the same reason, the delay performance has also significantly improved.

### 6.2. Multi-cell scenario

We now consider a cluster of three gNB–NR macro cells as shown in Fig. 12. We want to study the impact of inter-cell interference on the RAT selection algorithm running in a 5G-Flow network with this scenario. The inter-site distance for gNB–NR macro cells is 500 m. Each gNB–NR macro cell has multiple Wi-Fi APs. The distribution of UEs in the cells follows from the single cell scenario. We also consider four different service types as described earlier. The essential parameters relevant for this simulation scenario are the same as those of single cell scenario and given in Table 1. The RAT selection algorithm remains the same as given in Algorithm 1, except the channel condition for user $i$ under 5G-NR network ($ch_{g,i}$) also considers interference.

In this simulation scenario, we consider two cases from Table 2. Fig. 13 illustrates the result for Case b and Case c. It is important to observe that the overall network throughput in both the cases as compared to single cell scenario has decreased (Fig. 9) due to inter-cell interference. We now analyze the performance for each case. The performance of overall network throughput and average packet delay for Case b (Fig. 13a) under 5G-Flow network is better than that of standard 5G network. The performance of users requesting priority 2 service (a high priority service) experience significant degradation under standard 5G network, while the 5G-Flow network delivers the expected performance to these users. Note that the standard 5G network schedules the high-priority users to gNB–NR RAT. Due to interference, the capacity of gNB–NR is reduced; hence the performance of priority 2 users suffers under the standard 5G network.
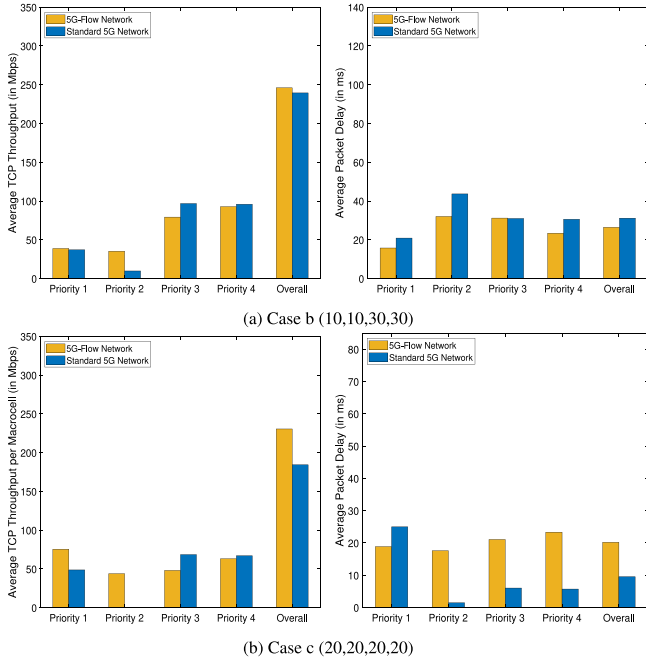
**Fig. 13.** Downlink performance of 5G-Flow network vis-a-vis 3GPP 5G network for a three cell scenario.

**Table 3**
MmWave simulation parameters.

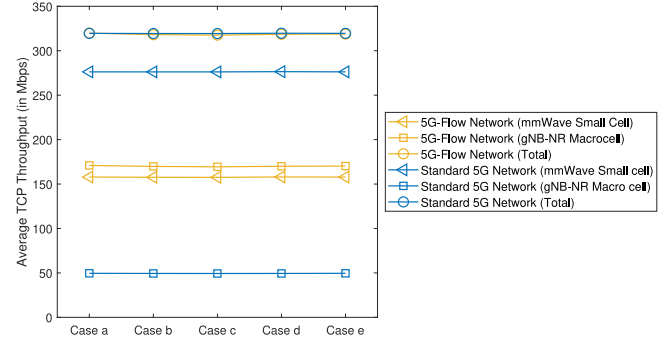| Parameters | Values |
| --- | --- |
| Carrier frequency | 28 GHz (TDD) |
| 5G numerology ($\mu$) | 2 |
| Bandwidth | 50 MHz |
| UE/mmWave transmit power | 30/23 dBm |
| UE/mmWave antenna gain | 2/26 dBi |
| UE/mmwave antenna height | 1.5/10 m |
| UE/mmWave noise figure | 7/10 dB |



**Fig. 14.** Load balancing across gNB–NR and mmWave cells.

The performance of Case c (Fig. 13b) under the 5G-Flow network as compared to the standard 5G network is much better. The observed TCP throughput for priority 2 is close to zero for Case c under the standard 5G network as the load of priority 1 users exhaust most of the available capacity of gNB–NR RAT. This does not happen for 5G-Flow network as the RAT is selected based on various parameters including the load on each RAT. Overall, we observe that the RAT selection algorithm under 5G-Flow network offers significant gains over standard 5G network under both single and multi-cell scenarios.

### 6.3. Single RAT heterogeneous network

In the above simulations, we have analyzed the performance of a heterogeneous network comprising a 5G NR based gNB–NR (macro layer) and Wi-Fi APs (micro layer). We now examine the performance of a single RAT heterogeneous network, where both the macro layer (macro cell) and the micro layer (small cell) use the same radio technology, the 5G-NR technology. In this setup, the Wi-Fi APs are replaced with NR millimeter wave (mmWave) small cells. There is no change in the macro cell configuration and the relevant parameters for this scenario (except the Wi-Fi AP network model) remain the same as given in Table 1. The simulation parameters considered for the mmWave cell are given in Table 3.

To simulate a heterogeneous 5G-Flow network with NR macro cell and mmWave (small) cells, we assume two separate radio interface stacks to control (support) macro cell and mmWave cells, respectively. These two separate radio interface stacks are mapped to separate interfaces on the MRN OpenFlow switch. For this simulation setup, we discuss the load balancing scenario across the macro and mmWave cells for different cases discussed in Table 2. The 5G-Flow network selects a cell for a UE between macro and mmWave cell using a similar algorithm given in Algorithm 1. The only change in the algorithm is the replacement of Wi-Fi AP related parameters with the mmWave parameters. A UE in standard 5G network selects a cell which offers a better SNR.

Fig. 14 compares load balancing for the given scenario under 5G-Flow with that of the standard 5G network. The overall performance is observed to be the same in 5G-Flow network and standard 5G network for all the cases. This is because both macro cell and mmWave cells use the NR radio interface which can render the required quality of service for all the cases, irrespective of the priority of the requested services. However, the main difference can be seen in the performance of individual cells. In Fig. 14, we observe that the capacity of gNB–NR macro cell in standard 5G network is underutilized as UE selects a cell based on SNR. In 5G-Flow network, it is the network that selects a cell for a UE and hence it can balance the load across macro and mmWave cells. Moreover, by moving a set of users from the mmWave cells (that are observing low load) to macro cell, some mmWave cells can be put into sleep mode. Therefore, the power consumption of the overall network may reduce considerably.

## 7. Impact on 3GPP 5G network architecture

It is important to note that the 5G-Flow architecture proposes solutions in areas that 3GPP 5G standard has left open for vendor implementation. For example, 3GPP has not defined specific interface between the RRC and NGAP stack on the gNB-CU or the interface between the NAS and the RRC layers on the UE. Our work achieves the unification of multiple RATs in the access network by standardizing such interfaces with the help of OpenFlow. Moreover, the proposed architecture is substantially software-based and can be easily implemented. Our proposal requires instantiation of OpenFlow switches at UE and MRN switch. Even though gNB-CU and N3IWF has been replaced with MRN OpenFlow switch and 5G-Flow Controller, the changes are more from the organization perspective. We use the existing networking concepts given by OpenFlow protocol such as logical ports to enable a uniform abstraction to be used by the 5G-Flow controller. The processing of RRC and underlying protocol stack or NGAP and underlying protocol stack is done as part of logical port processing. The 5G-Flow controller need not know the working of these protocol stacks, thereby making its implementation relatively simple.

The proposed software-related changes do not modify the existing protocols/interfaces of the 3GPP 5G architecture. The interfaces such as N2/N3 interface between RAN and 5GC, radio interface between UE and RAN, and F1 interface between gNB-CU and gNB-DU may continue to be used without any modifications under the 5G-Flow framework.

As 5G-Flow network proposes enhancements to the 3GPP 5G network, it is important to discuss security-related aspects. We envision 5G-Flow controller and MRN OF switches to be a part of an operator network, and therefore it should not introduce security-related vulnerabilities in the network. Moreover, MRN OpenFlow switch is placed beyond gNB–NR radio stack in the 5G RAN and does not disrupt the secure tunnel used over the radio interface between UE and gNB–NR radio stack. Similarly, NAS signaling exchanged between UE and 5GC is also encrypted. This is not disturbed by MRN OpenFlow switch (and 5G Flow Controller) as NAS messages are transparently passed through MRN OpenFlow switch.

Under the 5G-Flow network, data latency will be reduced in scenarios where a UE communicates directly with an external data network via RAN, bypassing the core. The signaling latency is also reduced as NAS signaling can be skipped. However, when a UE communicates with 5GC, additional signaling messages between UE OpenFlow switch and 5G-Flow controller may slightly increase the signaling latency. We view this as a reasonable trade-off required to decouple UE's communication with RAN from its communication with 5GC. Moreover, as there are only a few additional signaling messages, their impact on UE's latency and battery will be minimal.

## 8. Applications

In this section, we present some important applications of 5G-Flow RAN architecture below. Although, we discuss only two applications due to space constraints, 5G-Flow architecture can help support other use-cases too.

### 8.1. Direct Internet connectivity via RAN

A UE in the existing 3GPP cellular network can access Internet solely through the CN. It does not have the flexibility to connect to the Internet directly from the RAN bypassing the CN. This feature can be beneficial in those areas where users are mostly stationary, and therefore a data tunnel through CN may not be required. However, the advantages of the cellular stack with an efficient L2/L1 layer can still be exploited.

We now discuss the procedure to access the Internet directly bypassing the CN. A UE sends a packet to the IP interface of the UE OpenFlow switch and a table-miss is observed. On table-miss, the UE OpenFlow switch forwards the packet to the 5G-Flow controller via *OF Packet-in* message. Depending on the QoS requirement, the controller decides whether a dedicated DRB needs to be created. In case a dedicated DRB is required, the controller sends an *OF-Config* message to gNB–NR interface to create a logical port, which is translated by RRC layer and a DRB is created using *RRC Reconfiguration* messages. An OF-Config message is also sent to the UE so that it can create a logical port on the NR interface and map the newly established DRB to the logical port. The 5G-Flow controller adds the appropriate flow entries at UE and MRN OpenFlow switch so that the UE can access the Internet.

### 8.2. Simpler mechanism for 5G non-standalone deployment

It appears that 5G deployments will be carried out in phases, and the early adopters of 5G are most likely to choose a non-standalone deployment. This method involves deployment of eNB and gNB in RAN, which communicate with the 4G CN as shown in Fig. 15. A UE should support dual connectivity to both eNB and gNB to avail 5G services. Since gNB does not communicate with 4G CN, it is modified to communicate with 4G CN and referred to as en-gNB. An eNB acts as a master node while en-gNB acts as a secondary node, and they communicate via X2 interface. All signaling exchange between UE and the network happens via eNB, while gNB is only used for data transfer. The non-standalone method is a much faster way to deploy 5G as it
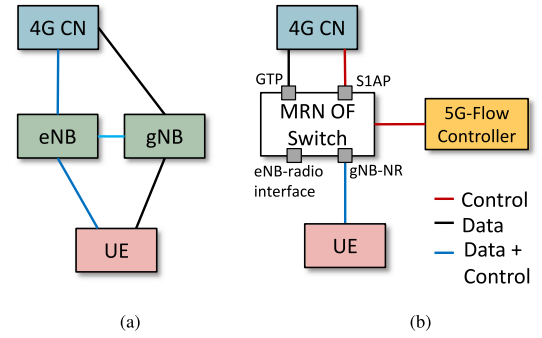


**Fig. 15.** Implementation of non-standalone 5G architecture proposed by (a) 3GPP, viz-a-viz (b) 5G-Flow Network.

leverages the existing 4G infrastructure. However, 5G RAN capability cannot be exploited entirely in the current architecture.

5G-Flow architecture enables non-standalone deployment in a much simpler manner. Due to the complete decoupling between RAN and CN connectivity, it allows a UE to use 5G NR interface to connect to 4G CN and receive mobile data services without necessarily using the dual connectivity.

## 9. Conclusions and discussions

Unified multi-RAT systems are indispensable in the next-generation networks. Although 3GPP 5G architecture supports multi-RAT integration at 5GC, there are several gaps in the existing 5G architecture. To address this, we re-architect the 3GPP 5G RAN to realize an integrated, software-defined multi-RAT RAN using OpenFlow Protocol. To realize the 5G-Flow RAN, we have suggested minimal software-based changes in the 3GPP 5G architecture which can be easily implemented. Moreover, we do not propose any changes in the protocol used between UE and gNB, UE and 5GC, and gNB and 5GC. The proposed architecture presents significant advantages over existing 3GPP 5G architecture such as (i) simplified multi-RAT integration, (ii) efficient dataflow management at RAN, (iii) flexibility to connect to any CN or Internet directly via 4G/5G/Wi-Fi based RAN, and (iv) simpler implementation of the non-standalone architecture. The performance evaluation of our architecture (using the simulator developed by us) shows promising gains over the 3GPP 5G network performance.

The current 3GPP 5G architecture is designed in a manner that drives user traffic through CN even when the user is stationary. In this paper, we present a method wherein a user can bypass CN and directly access Internet. This feature can be beneficial in effectively managing future networks that aim to serve billions of connected devices, of which a large fraction belongs to stationary IoT devices. Complete decoupling of radio and CN protocol stack in 5G-Flow RAN allows for greater flexibility to develop radio technology independent of the CN being used. Future networks can exploit this feature to facilitate faster development cycles for newer RATs.

In future, we aim to propose some enhancements in the OpenFlow and OF-Config protocols to support these 5G specific logical ports. The development of an open protocol for analyzing and translating NGAP messages for the 5G-Flow controller is an important future work. Additionally, we require a modified Yang model to be used with NETCONF, which can support event notifications based on the radio measurement data, NGAP messages, etc. We expect that mobility management can be effectively handled by introducing a virtual OpenFlow switch (at MRN OpenFlow switch) for every UE. We aim to examine these open issues in the future.

## CRediT authorship contribution statement

**Meghna Khaturia:** Conceptualization, Validation, Writing - original draft, Software. **Pranav Jha:** Conceptualization, Validation, Writing - original draft. **Abhay Karandikar:** Conceptualization, Writing - review & editing, Supervision, Project administration, Funding acquisition, Resources.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

## References

[1] 3GPP TS 23.501 v16.7.0, System Architecture for the 5G System, Technical specification, 2020, URL https://www.3gpp.org/DynaReport/23501.htm.

[2] 3GPP TS 24.193 v16.2.0, 5G System; Access Traffic Steering, Switching and Splitting (ATSSS), Technical specification, 2020, URL https://www.3gpp.org/DynaReport/24193.htm.

[3] OpenFlow Switch Specification, v1.3.0, Technical specification, Open Networking Foundation, 2014, URL https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-spec-v1.3.0.pdf.

[4] A. Gudipati, D. Perry, L.E. Li, S. Katti, SoftRAN: Software defined radio access network, in: Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN), 2013, pp. 25–30.

[5] T. Chen, H. Zhang, X. Chen, O. Tirkkonen, SoftMobile: Control evolution for future heterogeneous mobile networks, IEEE Wirel. Commun. 21 (6) (2014) 70–78.

[6] X. Foukas, N. Nikaein, M.M. Kassem, M.K. Marina, K. Kontovasilis, FlexRAN: A flexible and programmable platform for software-defined radio access networks, in: Conference on Emerging Networking EXperiments and Technologies (CoNEXT), 2016, pp. 427–441.

[7] H. Wang, S. Chen, H. Xu, M. Ai, Y. Shi, SoftNet: A software defined decentralized mobile network architecture toward 5G, IEEE Netw. 29 (2) (2015) 16–22.

[8] C.J. Bernardos, A. De La Oliva, P. Serrano, A. Banchs, L.M. Contreras, H. Jin, J.C. Zúñiga, An architecture for software defined wireless networking, IEEE Wirel. Commun. 21 (3) (2014) 52–61.

[9] I.F. Akyildiz, P. Wang, S.-C. Lin, SoftAir: A software defined networking architecture for 5G wireless systems, Comput. Netw. 85 (2015) 1–18.

[10] X. Jin, L.E. Li, L. Vanbever, J. Rexford, SoftCell: Scalable and flexible cellular core network architecture, in: Ninth ACM Conference on Emerging Networking Experiments and Technologies (CoNEXT), 2013, pp. 163–174.

[11] M. Bansal, J. Mehlman, S. Katti, P. Levis, Openradio: a programmable wireless dataplane, in: First Workshop on Hot Topics in Software Defined Networks (HotSDN), 2012, pp. 109–114.

[12] T. De Schepper, P. Bosch, E. Zeljković, F. Mahfoudhi, J. Haxhibeqiri, J. Hoebeke, J. Famaey, S. Latré, ORCHESTRA: Enabling inter-technology network management in heterogeneous wireless networks, IEEE Trans. Netw. Serv. Manag. 15 (4) (2018) 1733–1746.

[13] K.K. Yap, M. Kobayashi, R. Sherwood, T.Y. Huang, M. Chan, N. Handigol, N. McKeown, OpenRoads: Empowering research in mobile networks, SIGCOMM Comput. Commun. Rev. (CCR) 40 (1) (2010) 125–126, http://dx.doi.org/10.1145/1672308.1672331.

[14] K.K. Yap, M. Kobayashi, D. Underhill, S. Seetharaman, P. Kazemian, N. McKeown, The stanford openroads deployment, in: ACM International Workshop on Experimental Evaluation and Characterization, 2009, pp. 59–66.

[15] M. Yang, Y. Li, D. Jin, L. Su, S. Ma, L. Zeng, Openran: a software-defined ran architecture via virtualization, ACM SIGCOMM Comput. Commun. Rev. (CCR) 43 (4) (2013) 549–550.

[16] E. Coronado, S.N. Khan, R. Riggio, 5G-EmPOWER: A software-defined networking platform for 5G radio access networks, IEEE Trans. Netw. Serv. Manag. 16 (2) (2019) 715–728, http://dx.doi.org/10.1109/TNSM.2019.2908675.

[17] A.N. Manjeshwar, A. Roy, P. Jha, A. Karandikar, Control and management of multiple RATs in wireless networks: An SDN approach, in: IEEE 2nd 5G World Forum (5GWF), 2019, pp. 596–601.

[18] P.K. Taksande, P. Jha, A. Karandikar, P. Chaporkar, Open5G: A software-defined networking protocol for 5G multi-RAT wireless networks, in: IEEE Wireless Communications and Networking Conference Workshops (WCNCW), 2020, pp. 1–6.

[19] S. Chandrashekar, A. Maeder, C. Sartori, T. Höhne, B. Vejlgaard, D. Chandramouli, 5G multi-RAT multi-connectivity architecture, in: IEEE International Conference on Communications Workshops (ICC), 2016, pp. 180–186.

[20] D. Laselva, D. Lopez-Perez, M. Rinne, T. Henttonen, 3GPP LTE-WLAN aggregation technologies: Functionalities and performance comparison, IEEE Commun. Mag. 56 (3) (2018) 195–203.

[21] P. Nuggehalli, LTE-WLAN aggregation [industry perspectives], IEEE Wirel. Commun. 23 (4) (2016) 4–6.

[22] R. Enns, M. Bjorklund, J. Schoenwaelder, A. Bierman, RFC 6241 - Network Configuration Protocol (NETCONF), Technical specification, 2011, URL https://tools.ietf.org/html/rfc6241.

[23] OpenFlow Management and Configuration Protocol, v1.2, Technical specification, Open Networking Foundation, 2014, URL https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow-config/of-config-1.2.pdf.

[24] OpenFlow Notifications Framework, OpenFlow Management, v1.0, Technical specification, Open Networking Foundation, 2013, URL https://www.opennetworking.org/wp-content/uploads/2013/02/of-notifications-framework-1.0.pdf.

[25] 5G-flow simulator source code, 2021, URL https://github.com/MeghnaKhaturia/5G-Flow-RAN.

[26] 3GPP TR 38.901 v16.1.0, Study on Channel Model for Frequencies from 0.5 to 100 GHz, Technical report, 2020, URL https://www.3gpp.org/DynaReport/38.901.htm.

[27] IEEE 802.11-14/0882r4, IEEE 802.11ax Channel Model Document, Technical specification, 2014, URL https://mentor.ieee.org/802.11/dcn/14/11-14-0882-04-00ax-tgax-channel-model-document.docx.

[28] IEEE Std 802.11-2016, IEEE standard for information technology—Telecommunications and information exchange between systems local and metropolitan area networks—Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, 2016, pp. 1–3534.

[29] 3GPP TS 38.331 v16.3.1, 5G; NR; Radio Resource Control (RRC); Protocol Specification, Technical specification, 2021, URL https://www.3gpp.org/dynareport/38331.htm.

**Meghna Khaturia** has recently received her M.Tech and Ph.D. from Indian Institute of Technology Bombay. Her PhD work dealt with developing a Frugal 5G communication network for rural broadband access. Prior to that she did her B.Tech in Electronics and Communications engineering from Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat, India, in 2013. Her research interests include rural broadband, wireless networking, software defined networking, fog computing and network architecture.

**Pranav Jha** has extensive research and development experience in communication and networking technologies with expertise in network architecture, communication protocols and resource management algorithms. He currently works with IIT Bombay, Mumbai, India where he focuses on research programs under the broad umbrella of wireless networking with emphasis on SDN based architectures for Mobile Networks, Public Safety Communication and Rural Broadband Communication. He has multiple research papers, patents issued and pending, and contribution to IEEE and 3GPP standards.

**Abhay Karandikar** is currently the Director of IIT Kanpur (on leave from IIT Bombay). He served as Member (part-time) of Telecom Regulatory Authority of India (TRAI) from 2018 to 2020. In IIT Bombay, he served as Institute Chair Professor in the Electrical Engineering Department, the Dean (Faculty Affairs) from 2017 to 2018 and the Head of the Electrical Engineering Department from 2012 to 2015. Prof. Karandikar is the founding member of Telecom Standards Development Society, India (TSDSI), India's standards body for telecom. He was the Chairman of TSDSI from 2016 to 2018. His research interests include resource allocation in wireless networks, software defined networking, frugal 5G and rural broadband. Detailed biography can be found at https://www.ee.iitb.ac.in/~karandi/.