International Journal of Data and Network Science 7 (2023) 1421–1434

Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Improved flat mobile core network architecture for 5G mobile communication systems

Mohammad Hijjawi[a], Mohammad Al Shinwan[a*], Mahmoud H. Qutqut[a], Waleed Alomoush[b], Osama A. Khashan[c], Marah Alshdaifat[d], Abdullah Alsokkar[e] and Laith Abualigah[f,g,h]

[a]Faculty of Information Technology, Applied Science Private University, Amman 11931, Jordan
[b]School of Information Technology, Skyline University College, Sharjah P.O. Box 1797, United Arab Emirates
[c]Research and Innovation Centers, Rabdan Academy, Abu Dhabi P.O. Box 114646, United Arab Emirates
[d]Faculty of Engineering, Department of Civil Engineering, The Hashemite University, P.O. Box 330127, Zarqa 13133, Jordan
[e]Faculty of Business, Applied Science Private University, Amman 11931, Jordan
[f]Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19328, Jordan
[g]Faculty of Information Technology, Middle East University, Amman 11831, Jordan
[h]Prince Hussein Bin Abdullah College for Information Technology, Al Al-Bayt University, Mafraq, 130040, Jordan

C H R O N I C L E

Article history:
Received: December 2, 2022
Received in revised format: January 29, 2023
Accepted: March 25, 2023
Available online: March 25, 2023

Keywords:
Mobile core network
Mobility management systems
5G and beyond
4G network

A B S T R A C T

The current mobile network core is built based on a centralized architecture, including the S-GW and P-GW entities to serve as mobility anchors. Nevertheless, this architecture causes non-optimal routing and latency for control messages. In contrast, the fifth generation (5G) network will redesign the network service architecture to improve changeover management and deliver clients a better Quality-of-Experience (QoE). To enhance the design of the existing network, a distributed 5G core architecture is introduced in this study. The control and data planes are distinct, and the core network also combines IP functionality anchored in a multi-session gateway design. We also suggest a control node that will fully implement the control plane and result in a flat network design. Its architecture, therefore, improves data delivery, mobility, and attachment speed. The performance of the proposed architecture is validated by improved NS3 simulation to run several simulations, including attachment and inter- and intra-handover. According to experimental data, the suggested network is superior in terms of initial attachment, network delay, and changeover management.

1. Introduction

Recently, mobile data and traffic growth pushed mobile operators and service providers to re-engineer the core mobile network and deliver salable solutions through several solutions, such as flat mobile network architecture (Bruschi et al., 2019; Hu et al., 2020; Al Shinwan et al., 2021). It is anticipated that mobile data traffic will increase double per year in the upcoming years. For example, it's envisioned that Internet of Things (IoT) devices connections will exceed 100 billion in the upcoming years (Liu & Jiang, 2016; De La Oliva et al., 2015; Abualigah, Diabat, Sumari, et al., 2021). This creates several challenges for the network providers to build a mobile network architecture that can handle network data traffic. The 5g mobile network aims to provide a higher number of connected devices, and faster speed to overcome such challenges (Rost et al., 2017; Khan et al., 2020; Alsokkar et al., 2023).

The main challenges that need to be addressed are effectively handling data traffic passing via the core network and improving the network capacity (Uddin et al., 2021). Several solutions have been proposed to conquer the hierarchical and centralized

* Corresponding author.
E-mail address: m_shinwan@asu.edu.jo (M. Al Shinwan)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2023.3.021

core network to solve the data traffic. One of the solutions to flatten the core network is to decrease the number of network entities to deal with the number of connected devices such as IoT, and big data generators (Abd Elaziz et al., 2021; Abualigah, Diabat, & Abd Elaziz, 2021; Abualigah et al., 2023). On the other hand, different solutions such as radio resource management, small cell technology, and device-to-device communication can improve network capacity (Bouras & Diles, 2015; Qutqut, 2014).

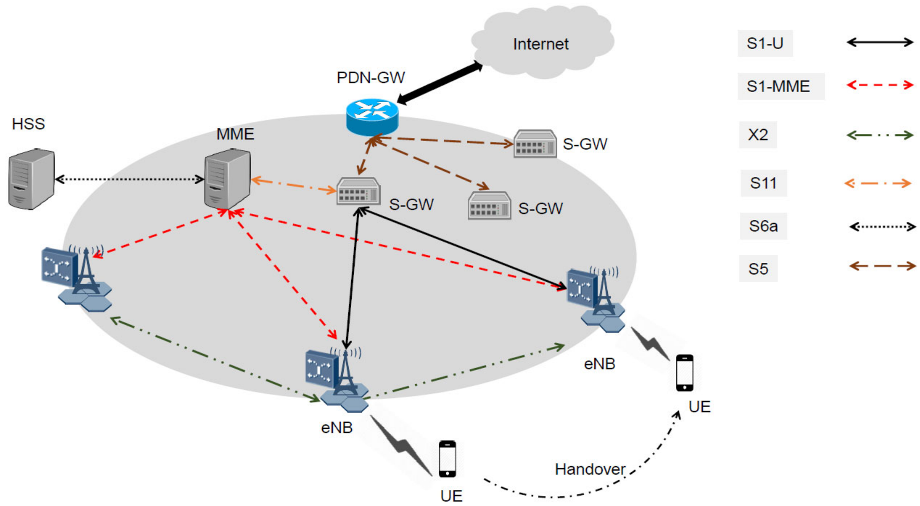

Fig. 1. The 4G mobile network Evolved Packer Core (EPC) (Al Shinwan & Chul-Soo, 2018).

Fig. 1 shows the 4G mobile network architecture. The main component in the 4G is Evolved Packet System (EPS), which is built based on a flat IP-based. The EPS network is divided into the Evolved Packet Core (EPC) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (Project, 2020). The EPC consists of Packet-data-network gateway (PGW), anchor mobility Serving Gateway (SGW), Home Subscription Server (HSS), and Management Mobility Entity (MME). On the other hand, the E-UTRAN includes the base station called eNodeB (eNB). Several bearers are used for the data routing between the User Equipment (UE) and the PGW, i.e., S1-U, S5/S8, and X2. The connection between the eNB and SGW is made by the S1-U. While traffic delivers between eNB and S-GW via S5/S8, and in the case of mobility, X2 is used for a temporary connection. The data plane consists of the PGW and SGW; the SGW plays the role of mobility anchor to deliver packets between the eNB and PGW. PGW serves as an edge router entity to allow the user to connect to the Internet, IP allocation, traffic routing, and packet filtering. The functionality of the centralization in the core network is managed and controlled by the MME. The GPRS tunneling protocol (GTP) managed the data delivery processes, including mobility and initial attack (Qureshi et al., 2020; Koubias & Haralabidis, 2001; Masa'deh et al., 2023).

Several approaches have been proposed to enhance the core network, specifically the EPC. One of the promising approaches is Software Defined Networking (SDN). Basta et al. (2013) and Nguyen et al. (2016) proposed an approach using OpenFlow protocol, based on classified network functions regarding their effect on the data plane and the control plane. This paper (Said et al., 2013), proposes a method utilizing the OpenFlow protocol through merging the controls protocols that operate on S11 and S1-MME. Most of the suggested methods are based on the SDN, which was created by separating the control and data planes by studying the OpenFlow protocol. These alternatives, however, are hampered by the centralized mobility design. Thus, to reduce network latency and disperse mobility, we introduce a distributed EPC network.

Recently the innovative technique known as distributed mobility management (DMM) has attracted academics and research. The DMM is divided into fully and partially architecture. The fully distributed network is distributed in both the data and control plane. Further, it's implemented in limited situations, such as metropolitan areas, and is complex to implement in global area networks. The control plane is centralized, despite the fact that the data plane is scattered over the network (Chan et al., 2011). The DMM is a promising approach to managing the massive number of users and internet flow in the mobile core network (Jeon et al., 2013; Ki et al., 2014). Giust et al. (2015) introduced two partially distributed schemes based on the Software-Defined Networking (SDN), namely SDN-DMM, and the other method is based on the PMIPv6, which modifies the classical PMIPv6 protocol. Furthermore, Cominardi et al. (2017) proposed a DMM based on PMIPv6 with an evaluation of DMM and SDN solutions.

Since the user's IP addresses are assigned and can be designated during mobility procedures, DMM solutions based on PMIPv6 demonstrate the previous.

Recently, different proposed approaches have been focused on a flat mobility network to enhance mobility and speed up the handover speed. In (Shinwan & Chul-Soo, 2017), an approach has been presented by removing the SGW entity and S5/S8 interfaces and the S2/S11 link directly to the PGW entity. In (Modeas et al., 2021), a novel approach proposes based on a couple of techniques to determine appreciative radio access. The first technique is put on the UE, while the second is put inside the network. This strategy seeks to meet users' price, energy consumption, service quality, and security needs. In this paper (M. Al Shinwan & Chul-Soo, 2018), For mobility management and data transfer in the core mobile network, the IP-in-IP protocol is recommended as an alternative to the GTP protocol. Another strategy proposes combining the SGW and PGW to operate as a single entity. The new entity is dispersed throughout the mobile core network to govern and control mobility. In addition, the Mobility Management Entity (MME) allocates the IP address for the UE and functions as a mobility anchor (Al Shinwan et al., 2017; Al Shinwan et al., 2022).

The 3gpp technical standardization reports illustrated different solutions for dividing the control plane (CP) and user plane (UP) for the EPC network. Therefore, we propose to keep the CP acting as it's while the UP functionality will be distributed across the EPC. During the PDN establishment, the UP functionality will be distributed. Thus, the relocation and modifying of the session identification are determined by UE mobility. The extra signaling will stop if this strategy is used. as described in (Project, 2020; Project, 2016; Project, 2021).

In mobile networks, handover typically refers to a user's travel inside the network to maintain their current session. It has been noticed that calls frequently drop as a result of latency (Alsarhan et al., 2023). Gupta et al. (2021) proposed a smooth handover solution based on applying a fuzzy-based decision strategy for the mobile user utilizing a predictive manner to address this issue. In order to choose the next target eNodeB and determine the best handover time, the Multi-layer Feed Forward Network algorithm is suggested. The Access Network Selection Function (ANSF) is additionally utilized to choose the best access network currently accessible. The ANSF is used in conjunction with a number of inputs, including battery life, signal strength, data dependability, and mobility, to determine the best network layout. This paper (Sumathi et al., 2022) introduces a novel algorithm to improve mobile handover management in the device-to-device network and reduce the impact of the ping-pong effect. The handover management algorithm is based on the Reference Point Group Mobility (RPGM) model. Moreover, to determine the proper method of communication in networks, the mode selection method is implemented. Different metrics validate the proposed approach, such as the throughput, excessive handover rate, and absent handover speed rate. The simulation results demonstrate significant enhancement in the decreased absent handover speed rate and stable throughput. The simulation results demonstrate significant enhancement in the decreased absent handover rate and stable throughput. Different optimization algorithms that can be implemented to improve the handover process are found in (Abualigah, Shehab, et al., 2021; Abualigah et al., 2022; Abualigah et al., 2022).

For handover management, several approaches have been proposed. Kosmopoulos et al. (2022) used the capabilities of the Media-Independent Handover (MIH) and Fast Proxy Mobile IPv6 (FPMIP) protocols to perform handover in 5G vehicle networks. The scheme supports both proactive and reactive handover scenarios. Each vehicle is prepared for both handover situations using a velocity and alternate network monitoring technique. Sun et al. (2020) suggested using the coordinated multipoint (CoMP) protocol as a transmission method. By measuring dwell duration, the method seeks to intelligently follow user movement to allocated small or microcells. In order to increase dependability, the authors also propose a movement-aware CoMP handover improvement termed iMACH. The numerical and simulated findings demonstrate superior performance in terms of throughput, coverage, and handover probability. Bilen et al. (2017) suggest a mobility and resource estimation technique based on SDN to address the handover latency issue. The Markov chain approach is utilized to estimate the mobile node's neighbor eNB transition probabilities as well as their likelihood of having resources available. The best eNBs can then be virtually assigned to mobile nodes using an elegant mathematical framework, and all connections can be made using OpenFlow tables. The proposed approach is compared with the LTE network, and the results show that the suggested technique cuts down on the handover delay.

In this study, a core mobile network based on SDN architecture is introduced for use in a 5G communication system. We suggest decentralizing the control plane functionality with the SDN and distributing the user data plane. By reducing the number of network entities in the mobile core and simplifying the signaling messages, this strategy aims to lower overall data latency. Our analysis focuses on the necessary data plane operations, such as data delivery, initial attachment, and intra- and inter-gateway handover. To validate the effectiveness of the newly implemented network architecture, experiments are carried out using NS-3 simulation in a variety of scenarios. The SGW, PGW, and MME are now represented by a new network object known as the Mobile Control Unit (MCU). The core network comprises many control commodities to prevent a single point of defeat. The Mobility Packet Gateway (MPG) serves as both a data packet router for the internet and a mobility anchor, and the appropriate interface incorporates the CP and DP.

The rest of this paper is organized as follows; Section 2 introduces and describes the suggested 5G network. The specifics of how the new procedures operate are provided in Section 3. Section 4 provides the numerical analysis. The simulation setting and results are described in Section 5. Finally, Section 6 concludes this paper.

2. Proposed network architecture

The suggested 5G network infrastructure is depicted in Figure 2 in order to satisfy the flat distributed concept's requirements, the control protocols that operated on SGW and PGW to MME were removed (Hakeem et al., 2020). MME, SGW, and PGW were combined to form the Mobility Control Unit (MCU).

The MCU concentrates on management as well as control of the EPC core.

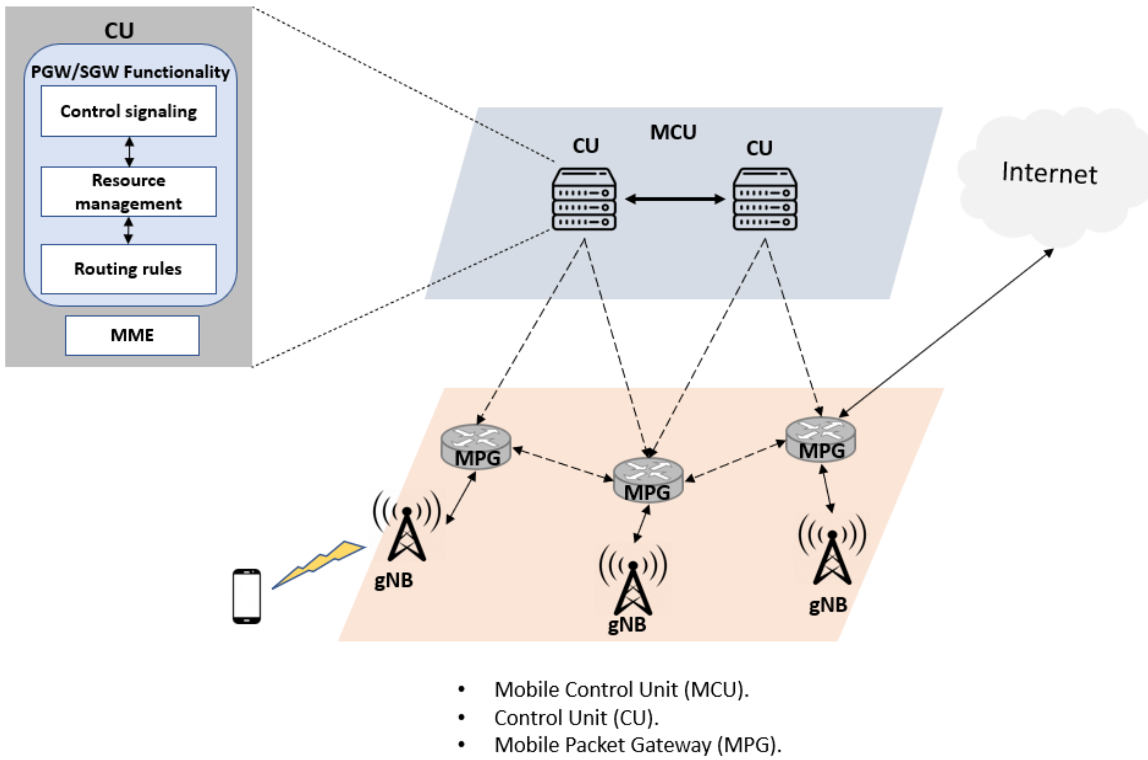

Fig. 2. Proposed distributed flat 5G core architecture

MCU acts as a central control unit responsible for responding to the interconnected functions and triggers a control signaling message. In addition, the MCU is responsible for the following function:

Allocate the UE IP address.

- Data forwarding.
- Establish DP session through assigning several tunnel parameters (such as Guaranteed Bit Rate (GBR), Quality Class Identifier (QCI), and Tunnel Endpoint Identifier (TEID))
- UE authentication and authorization.
- Track UE location.
- Bearer control and gateway selection.

The software-defined networking aims to separate CP and DP (Foundation, 2012). Further, to prevent a single network entity failure, MCU is divided into several Control Unites (CU's). Each CU oversees a region serviced by a collection of Gateways and E-UTRAN. Therefore, while the UE moves within the same CU domain, the UE IP address remains the same. Nevertheless, when the UE switches to a different CU domain, the CU provides a new IP address.

In a 4G network, the Mobility Management Entity (MME) is in charge of monitoring area updates, paging operations, and TAU. Therefore, with the growth of IoT devices and UE (Al-Sakran et al., 2018), the signaling traffic will be massive, hindering MME performance. We present a flat distributed DP that makes use of a larger gateway number and is based on non-anchor mobility. Therefore, the latency is reduced since the packet is routed to the closest Mobility Packet Gateway (MPG) node. In this work, we propose one MPG responsible for routing the IP packets to the Internet and acting as a mobility anchor. CU allocates the MPG based on the distance between the current UE location and the attached MPG.

Furthermore, the MPG's function is to receive DP routing rules from CE then attach them to the corresponding nodes. Also, the MPG manages the DP classification and filtering to distinguish the traffic regarding the policy for uplink and downlink and user profile. Finally, the GTP protocol is kept as it's used in the current 4g core network.

The CP will act as a centralized mobility anchor and represent the SDN in the proposed network. Therefore, the architecture of the current mobile network will not change in structure without adding more devices to support the SDN technology. In addition, the CP number expanded and spread in the core network and supported the distributed DP and data routing.

3. The procedure of the proposed architecture.

3.1 Initial attachment

In the 4G network, when a user first turns on his phone in the 4G network, an initial attach takes place to register the user with the mobile network and establish a connection to the Packet Data Network. (PDN) (*Netmanias Technical Document: EMM Procedure 1. Initial Attach: Part 2. Call Flow of Initial Attach*, 2014). The following steps describe the 4G attachment procedure:

- UE creates a communication with the eNB as data delivery and synchronization procedure.
- eNB sends these communication messages to update the location of the UE through the MME and Home Subscriber Server (HSS).
- The MME informs the SGW to establish a transmission path.
- SGW modified the bearer with PGW and accepted the UE attachment through the MME.
- The UE transmits an attachment procedure to the MME to modify the bearer via the SGW.

However, in the proposed network, the attachment procedure is simplified as follow:

- The CU provides the IP address and the registration data to the UE. As a result, the UE must travel to the CU to get the relevant data, which shortens the attachment time.
- When the UE needs to send data packets, the MGW will create a data path between the UE and GW.

Fig. 3 illustrates the attached technique in detail and provides the following details:

The UE sends an Attach Request message to the gNB that contains the International Mobile Subscriber Identity (IMSI). The gNB then uses the Initial UE context message to deliver the Attach Request message to the CU after adding additional parameters like the E-UTRAN Cell Global Identifier (ECGI) and Tracking Area Identifier (TAI). It is assumed that the procedures for authentication, authorization and location updates are unchanged. After completing these steps, CU will issue an IP address to UE and choose the closest getaway from its IP pool. To build an uplink GTP tunnel, the gNB gets an Attach accepts message, including MPG information such as UE IP, gateway IP, UL-TEID, and QoS. The radio bearer is set up to use UE, and the gNB changes the radio bearer's state from de-registered to registered. A session control message and gNB information are transmitted by CU to establish a downlink GTP connection (IP gNB, DL TEID, QoS). Trigger messages are issued, and up/down links GTP tunnels are built between the eNB and gateway if UE requires an instant Internet connection.

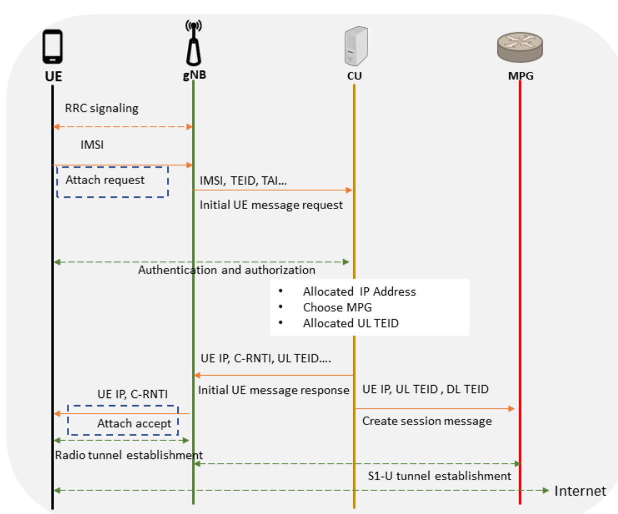

Fig. 3. The 5G network model's proposed initial connect procedure.

3.2 Data delivery process

Depending on the destination address, the UE can transmit and accept packets across the GTP-tunnel in the mobile network between the MPG and gNB. Mobile node to mobile node and mobile node to the internet are the two types of data transmission approaches.

3.2.1 Mobile node to Internet

Fig. 4 shows the data flow between the mobile host and the internet in the planned 5g network. The UE transfers the packet to the closest gNB, which will then use the GTP tunnel protocol to encapsulate it and send it to the MPG. The MPG then determines the ideal route for forwarding the data packet to the internet. Therefore, there is no requirement for the SGW and PGW to exchange control messages. Additionally, this structure improves signaling, shortens transmission times, and makes tunnel formation for S5/S8 easier.

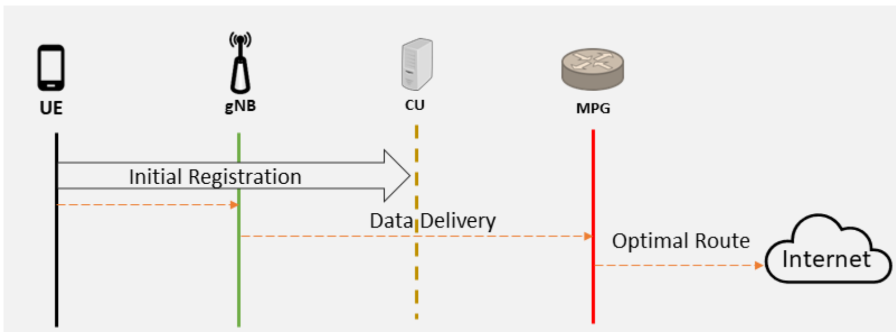

Fig. 4. The method used to deliver data from a mobile node to the internet.

3.2.2 Mobile node to Mobile node

The data delivery between two mobile hosts is illustrated in Fig. 5—the communication starts when the UE sends the data to the MPG through the GTP tunnel. To determine the destination address, the MPG entity decrypts the packet. The MPG entity sends an info request message to the CU entity to retrieve the MPG_{cn} IP address when it decides that the destination IP address belongs to another network. Next, MPG selects the optimal path to the gateway according to the information acquired from the CU. A paging mechanism takes place to locate the UE after the MPG_{cn} receives the packet. The GTP tunnel is then established to deliver the packets to the CU.

Fig. 5 shows how data is delivered between two mobile hosts. Communication begins when the UE delivers data to the MPG via the GTP tunnel. To determine the destination address, the MPG entity decrypts the packet. The MPG entity sends an info request message to the CU entity to retrieve the MPG_{cn} IP address when it decides that the destination IP address belongs to another network. Using the data obtained from the CU, MPG then chooses the best route to the gateway. A paging mechanism takes place to locate the UE after the MPG_{cn} receives the packet. The GTP tunnel is then created to deliver the packets to the CE.

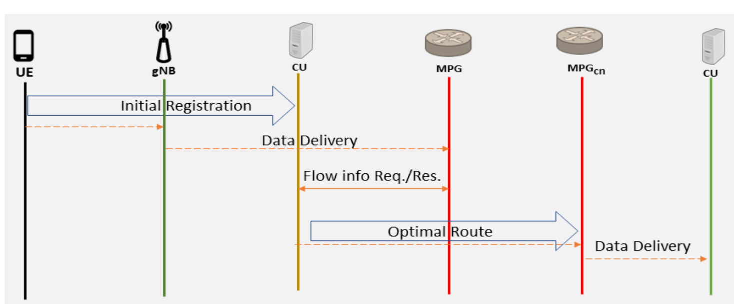

Fig. 5. Mobile node to Mobile node data delivery process.

3.3 The procedure of Intra-Gateway Handover

Handover (HO) is an essential process to maintain UE connectivity while moving inside the E-UTRAN area. Inter gateway handover (Inter-HO) and Intra gateway (Inter-HO) handover are two categories for the HO. When a user switches from a source gNB to a target gNB in a different location, the Inter-HO takes place. However, the Intra-HO occurs when the UE transfers from the source gNB (gNB-S) to the destination eNB (eNB-T) in the same domain area. Due to the fact that in the proposed work, E-UTRAN procedures and functions are unaffected, the Intra-HO with X2 asset scenario is identical to the 4G handover process (*Netmanias Technical Document: EMM Procedure 6. Handover without TAU: Part 2. X2 Handover*, 2014). Fig. 6 depicts the recently suggested handover without a gateway relocation process to lessen signaling messages.

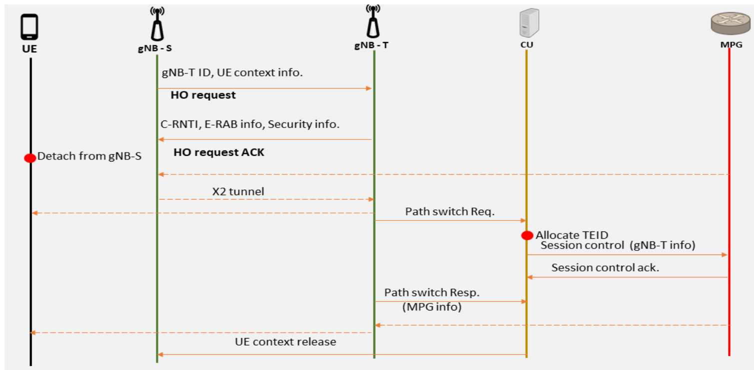

Fig. 6. The procedure of the introduced 5G network handover without gateway relocation

The CU entity manages all control operations in the proposed 5G network architecture; eliminating the MME entity will streamline the handover processes. The handover process goes as follows: the gNB-T makes a path switch request to the CU, asking it to allocate the necessary gateway information, add gNB-T parameters, and transfer the data to the MPG via a message session control. The Path switch response is then transmitted to the gNB-T. The GTP protocol is then constructed to transmit packets between the MPG and gNB-T.

3.4 The procedure of Inter-Gateway Handover

The Inter-HO is detected when the UE moves to a different gateway area. Fig. 7 shows the 4g architecture and the proposed architecture, respectively. The signaling messages during the handover procedure were reduced significantly by removing the role of PGW. Further, the proposed MGWs are distributed in the core network, allowing direct communication between the different MGWs. Fig. 8 illustrates the handover procedure of the proposed work, including establishing inter M-GW tunnel, X2 tunnel, and routing to the destination gateway (GW).

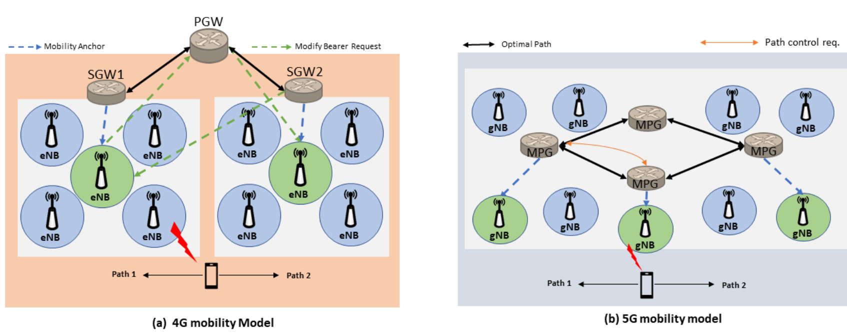

Fig. 7. The mobility with gateway relocation in (a) 4G network and (b) introduces 5G network.

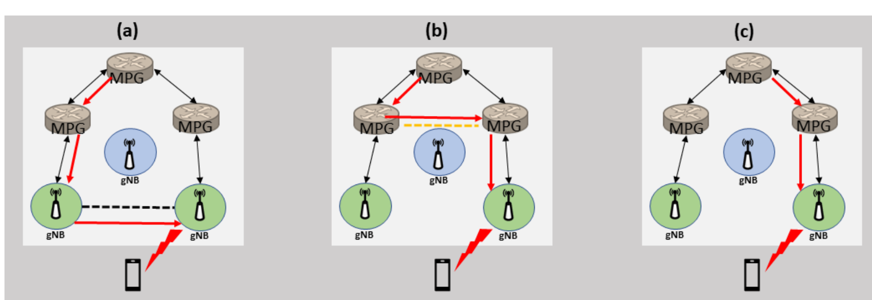

Fig. 8. Three essential steps to the handover process with gateway relocation in the proposed work: (a) setting up an X2 tunnel, (b) creating an inter-MPG tunnel, and (c) switching the destination gateway.

The X2 tunnel's construction is equivalent to the current 4G architecture. The gNB-S decides to HO and chooses the gNB-T. After transmitting the required messages to other gNBs, the handover takes place depending on the signal strength and neighbor cell index information. The HO request is made to the gNB-T while the X2 signaling is being prepared. Next, an X2 tunnel is established with gNB-T to send the data once the gNB-T replies with an acknowledgment message indicating that it can service the UE. The process is then starting to be detached from the gNB-S and attached to the gNB-T by UE.

In this work, inter gateway tunneling (IGT) is proposed to send the messages directly to the gateway. In the 4g network, the GTP tunnel was established by SGW to the PGW via sending modified bearer requests. However, in the suggested network architecture, packets are sent directly from the IGT to the source and target MPG (MGW-S and MGW-T). When the UE travels to another gNB belonging to a new MPG, the gNB-T sends a path switch request to the CU entity based on the MPG. Then, the CU assigns the needed parameters and distributes the DP to the gNB and gateway. For the tunneling between the core network devices and the E-UTRAN, the GTP protocol is used. Finally, after creating the IGW, the packet route is the best path via a direct tunnel, thus improving the performance and reducing the end-to-end latency.

The final phase is to route the data to the MPG-T. In this case, the packets are initially routed from the MPG_n located close to the MPG-S. The MPG_n has two cases: the first is when the MPG_n under the current CU, CU re-route the data path via request message. Second, the MPG_n belongs to a different CU; in this case, the CU requests information from the MCU entity, then the data is routed to the data path.

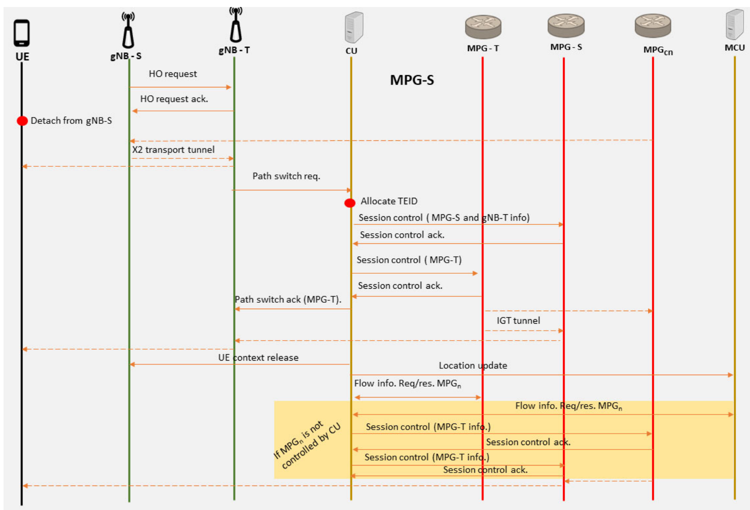

Fig. 9. The proposed 5G architecture's handover with gateway relocation.

The handover with gateway relocation illustrates in Figure 9. The X2 tunnel is created in this phase, and the gNB-T sends a request to switch the path to CU. CU decided that the MPG is relocating to be MPG-T. Then CU sends to MPG-S and MPG-T a session control message. Next, CU sends a path switch message to reply from MPG-T to the gNB-T. The gNB-T terminates the X2 tunnel and establishes the IGT tunnel and GTP tunnel to the MPG-T. The CU sends the UE context release to notify the gNB-S that the handover succeeds and eliminates the resources. Finally, the CU requests MCU or MPG by transmitting a flow info request asking for information about the MPG_n. Next, the CU forwards information about MPG-To MPG using a session control message. Then, a tunnel is established to re-route the packets between the MPG and MPG_n.

4. Mathematical Analysis

Here we explain the proposed architecture and the mathematical model used to determine the transmission latency for the 4G network. We assume that a specified message of size L is sent between two nodes through a wired link while disregarding wireless link latency.

The cost signaling of the first attached delays and HO is shown in the following expressions.

We express $\Phi_{A,B}\,(L, N_{A\to B})$ to represent the transmission delay of a message of size L sent over a wired link from A to B, where $N_{A\to B}$ denotes the number of wired hops. The following is how the transmission delay is indicated:

$$\Phi_{A,B}\,(L, N_{A\to B}) = L_{A\to B}\left[\frac{L}{BW} + DL + QD\right] \tag{1}$$

The default parameter and notations used in the performance analysis are described in Table 1.

The proposed work's initial attachment is outlined in section 3. Two different sorts of messages are required for this procedure: the control messages $C^{cn}_{Initial}$ and data massages $\Phi^{dm}_{Initial}$. Accordingly, the following is how the overall transmission delay for the initial attach is expressed:

$$\sum \Phi^{cm}_{Initial} + \Phi^{dm}_{Initial} \tag{2}$$

The suggested architecture's control messages can be summed up as follows:

$$\Phi^{cm,5G}_{Initial} = 2\Phi_{\Psi} + 2\Phi_{\omega} \tag{3}$$

and the data message

$$\Phi^{dm,5G}_{Initial} = 2\Phi_{\eta} \tag{4}$$

In the 4G network, the control and data messages for the transmission delay can be donated in Eq. (5) and Eq. (6):

$$\Phi^{cm,4G}_{Initial} = 5\Phi_{\Psi} + 2\Phi_{\Pi} + 4\Phi_{\Theta} + 2\Phi_{\omega} \tag{5}$$

$$\Phi^{dm,4G}_{Initial} = 2\Phi_{\zeta} + 2\Phi_{\omega} \tag{6}$$

As previously mentioned, the handover delay is divided into the Intra (Φ^{HO}_{Intra}) and Inter (Φ^{HO}_{Inter}). For both the proposed network and the 4G network, the intra handover delay can be estimated as follows:

$$\Phi^{HO,5G}_{Intra} = 2\Phi_{\zeta} + 3\Phi_{\Psi} + 2\Phi_{\Theta} + 2\Phi_{\omega} \tag{7}$$

$$\Phi^{HO,4G}_{Intra} = 2\Phi_{\Pi} + 2\Phi_{\Psi} + 5\Phi_{\Theta} + \Phi_{\omega} + \Phi_{\eta} \tag{8}$$

and the Inter handover latency is given in Eq. (9) and Eq. (10).

$$\Phi^{DL,5G}_{HO,Inter} = 3\Phi_{\zeta} + 2\Phi_{\eta} + 2\Phi_{\omega} \tag{9}$$

$$\Phi^{DL,4G}_{HO,Inter} = 6\Phi_{\Psi} + 2\Phi_{\Theta} + 2\Phi_{\omega} \tag{10}$$

Table 1
Default weights and values are utilized in the numerical analysis

Parameters	Description	Value
DL	Delay	2ms
QD	Queuing delay	2 ms
cm	Control messages size	50 bytes
dm	Data messages size	200 bytes
ζ	Hop counting between gNBs	2
Θ	Hop counting between CU and MPG	2
Π	Hop counting between HSS and CU	3
Ψ	Hop counting between gNB and CU	2
ω	Hop counting between SGW and PGW	3
η	Hop counting between gNB and SGW	2

5. Experimental results

5.1 Numerical results

Using the mathematical equations provided in Section 4, we adjust the variables indicated in Table 1, to compare the performance of the proposed architecture and the legacy network structure.

Fig. 10 shows how queuing delay (QD) affects the total transmission latency at each network node. The 4G network, completely distributed, and proposed techniques are contrasted in this graph (partially distributed method). For all networks, the delay grew linearly. The outcomes demonstrate that the suggested methodology produces competitive performance. This outcome is the consequence of the data packets in the 5G network being transmitted along the best route; in contrast, the data route in the 4G is provided by PGW and SGW.

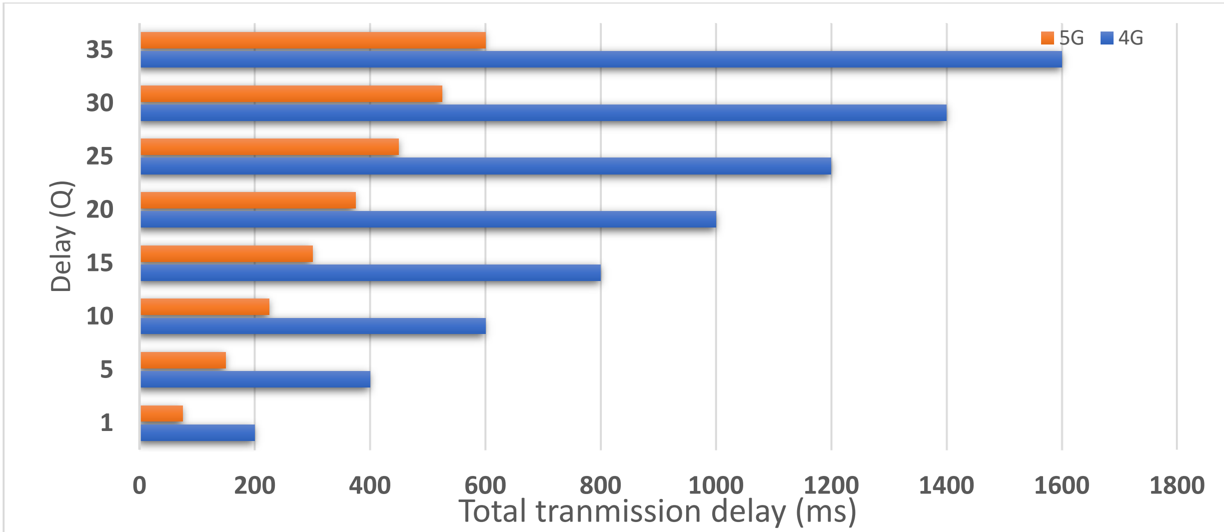

Fig. 10. The queuing delay (QD) impact at each node.

By changing Ψ, Fig. 11 illustrates how the hop count between the gNB and CU affects the transmission latency. The proposed architecture outperforms the 4G network in terms of performance because data packets are routed between the MPG and gNB optimally rather than using a centralized anchor between the SGW and PGW as in 4G.

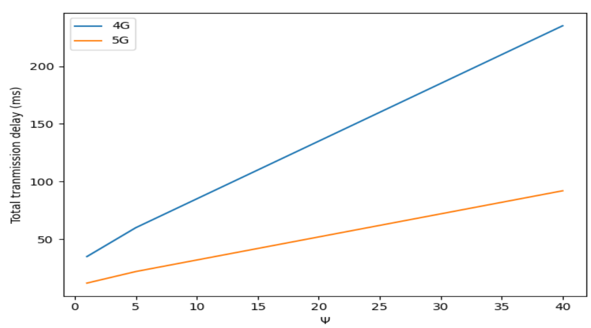

Fig. 11. The effect of changing Ψ on the overall transmission delay.

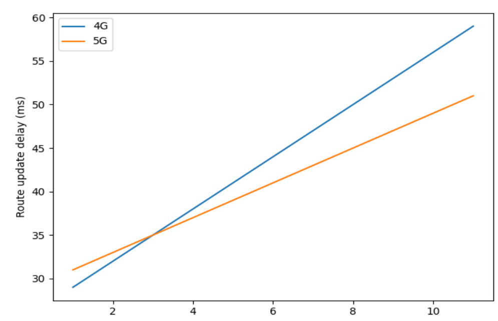

Fig. 12. The impact of ψ on inter-gateway handover.

In both networks, Fig. 12 compares intra-gateway handover. The outcome demonstrates how hop count affects gNB and CU (Ψ). Variable Ψ had a substantial impact on the 4G network's handover delay. However, because gNB works directly with CU to accomplish the routing, the suggested network shows a significant improvement. In contrast, In the 4G network, handover is implemented between the PGW and MME. A delay in inter-gateway handover is seen in Fig. 13. The outcome compares the handover delays between the identical network ω eNBs. Because the MME switches the bearer via SGW and PGW as a centralized anchor, the impact of varying *ω* significantly influences the 4G network delay. Nevertheless, because the CU executes path routing in conjunction with the MPG, the 5G architecture is unaffected by modifying *ω*.

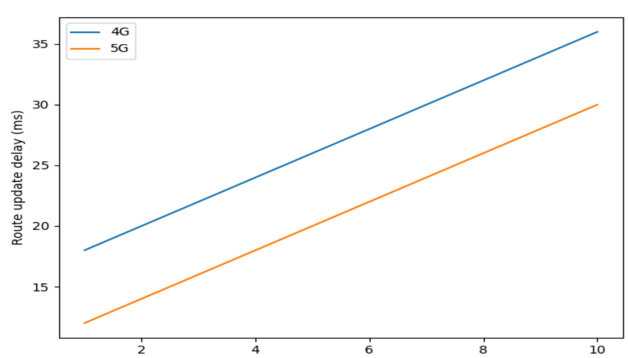

Fig. 13. The impact of ω on the handover delay.

5.2 Simulation modeling

In this part, we use the NS3 simulator (LENA simulation) to assess the performance of the suggested model. The two main elements of the 4G network topology that NS-3 has already created are the LTE and EPC models (Baldo et al., 2011).

The LTE framework is broken down into two categories: the higher radio stack, which contains the RLC, RRC, and PDCP protocols, and the lower radio stack, which contains the MAC and PHY layer, and Scheduler. The presented 5G network model keeps the access layer of 4G network; hence these components' functionality within the UE and eNB nodes remains unaltered. The EPC framework has been modified as follows.

- The SGW and PGW function as a single network entity, and the S5/S8 interface and gateway relocating mobility are not supported.
- Replacing the tunneling protocol GTPv1 with GTPv2.
- The socket transmission is used for the data plane, and logical connections are operated to correct erroneous results.

Hence, using a single SGW-PGW called MPG, as mentioned previously, results in improper signaling message protocols after running the simulation. To address these issues, we add a brand-new node that serves as the SGW, to which we attach the mobility features of the first SGW/PGW node. The S5/S8 GTP connection is then created between SGW and PGW, and complete signaling message implementation follows. The 4G simulated network mode now provides more realistic results for both the CP and DP.

The presented network architecture was built using LENA NS-3 version 3.22 in a Linux environment. The initial step is the construction of MCU, CU, and MPG nodes. The second step includes the control and data plane into the relevant interface. This work focuses on inter and intra handover gateway mobility simulations. While UE moves at various speeds and traverses multiple domains, the results are contrasted in terms of delay. Data from the simulated network is collected and stored using the flow monitor module, (Carneiro et al., 2009). A remote host node that we developed can transfer packets to other nodes and serves as an Internet server. The simulation model's additional parameters are displayed in Table 2 (Makaya & Pierre, 2008).

Table 2
Key parameters for simulation.

Parameter	Setting
UE speed	5 to 120 km/h
gNB Tx Power	46 dBm
Distance between gNB	100m
EPS Bearer type	NGBR-VIDEO-TCP
QCI	9
Bandwidth	5 MHz
Data rate	100 ps

5.2.1 Simulation results and discussion

UE must make an initial attack and create a GTP tunnel from the gNB to the MPG at the start of our scenario. The steps for re-attachment during the handover procedure resemble those for attachment. In the revised attachment scenario, we streamline the signaling process between the gNB and the gateway, resulting in a shorter average total time for the proposed 5G. In the event of a handover, this results in a shorter total time for re-attachment. Fig. 14 shows the total initial attachment times for 4G and the projected 5G.

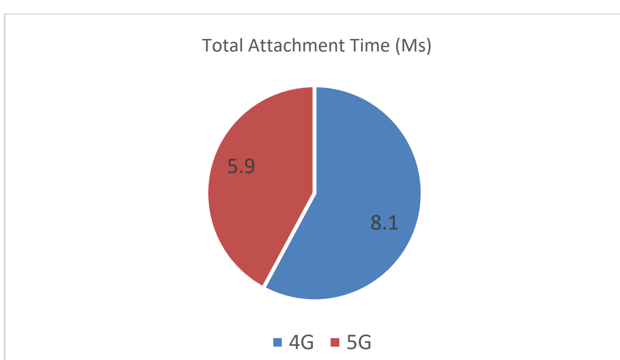

Fig. 14. Proposed 5G and 4G initial attachment latency.

5.2.2 Intra-Gateway mobility with X2 handover

Intra-HO happens when UE enters a new domain outside the present base station's coverage region, as mentioned in the handover section above. In the first simulation scenario, UE travels between the source and target gNB at speeds between 18 and 120 km/h. The time required for attachment/reattachment and data delivery makes up the entire execution time. The eNBs are 60 meters apart. The flow monitor module retrieves the value delaySum when the handover operation has successfully concluded. Fig. 15 indicates that for various UE speeds, the proposed 5G architecture reduces total delay time by 40% compared to the legacy 4G architecture. This illustrates how the suggested framework can save time and produce better results.

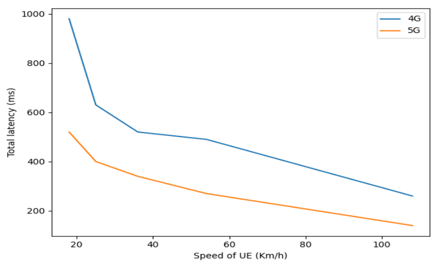

Fig. 15. Comparison of 5G with the 4G in terms of intra-HO latency for different UE speeds.

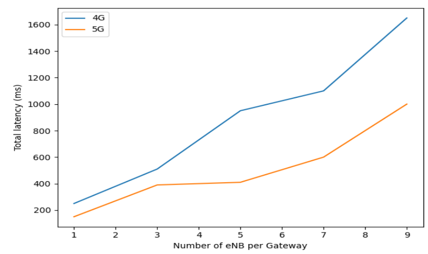

Fig. 16. The performance of 5G with the 4G networks in terms of latency of Intra-HO with gNB per Gateway.

The second simulation senior examines the performance when UE travels between various gNB inside the same gateway's domain. The number of gNBs nodes changed Between 2 and 6. The gNBs are 60 meters apart. All gNBs connect to a single SGW/PGW and the proposed single MPG. UE. travels over the territory of gNB at a speed of 108 km/h. The value delaySum is gathered from the flow monitor after 2 seconds have passed since UE re-attached to different gNB. The outcome is shown in Fig. 16. As anticipated, the total delay time is decreased by more than 30% using the suggested 5G approach.

5.2.3 Inter-Gateway mobility with X2 handover

The Inter-Gateway mobility is carried out using both the interior gateways tunnels and the X2 tunnel. In this simulation, several MPGs are created and connected to a single gNB. In the first simulated scenario, UE moves at a variable speed between the domains of two separate gateways. The result is shown in Fig. 17. The second simulated scenario shows inter-gateway mobility as the UE continuously moves across the gateway's domains. The proposed 5G with a distributed gateway, as shown in Fig. 18, reduces non-optimal path concerns and enhances transmission performance, in contrast to 4G, which necessitates data traffic termination and re-routing at PGW. The 5G network decreases total delay time by 14% in the first simulation with a different scenario and by 28% in the second simulated scenario when compared to the 4G network model.

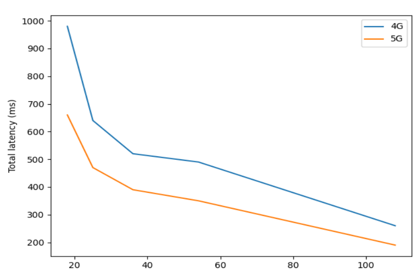

Fig. 17. Results of 5G and 4G network simulation with varied UE speed and latency of Inter-Handover.

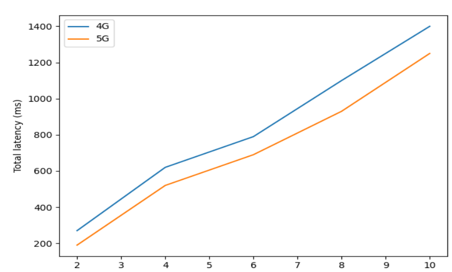

Fig. 18. 5G and 4G performance in terms of latency of Inter-Handover with gNB per Gateway.

6. Conclusion

This study offers a distributed user-plan architecture flat 5G core network to improve routing, lower latency, and signal across various network elements. Removing the centralized anchor nodes and gateways, which affect the initial attack time and the data channel transfer, results in a novel partially dispersed mobility for 5G and beyond networks. We have added a number of new entities to the simulator and merged the functionality of a number of existing ones based on NS-3. We demonstrated that our solution complies with 5G standards and 3GPP technical reports. To verify the effectiveness of the suggested design, experiments utilizing NS-3 simulation are carried out in various scenarios. To prevent a single point of failure, we first create new network entities that reflect the mobile core, like the MPG, which combines SGW, PGW, and MME. A new Mobility Gateway also serves as the network's anchor mobility while directing IP packets to the Internet. After that, add the control and data planes to the appropriate interface. The proposed design performed better for overall transmission latency, inter- and intra-handover delay, queuing delay, and total attachment time. The numerical analysis and simulation results showed that the suggested architecture design outperformed the old architecture by reducing end-to-end latency. Moreover, UE speeds up handover processes and switches between several gateways. The described architecture transmits data packets via the best possible routing paths.

References

Abd Elaziz, M., Abualigah, L., & Attiya, I. (2021). Advanced optimization technique for scheduling IoT tasks in cloud-fog computing environments. *Future Generation Computer Systems*, *124*, 142-154.

Abualigah, L., Diabat, A., & Elaziz, M. A. (2021). Intelligent workflow scheduling for Big Data applications in IoT cloud computing environments. *Cluster Computing*, *24*(4), 2957-2976.

Abualigah, L., Diabat, A., Sumari, P., & Gandomi, A. H. (2021). Applications, deployments, and integration of internet of drones (iod): a review. *IEEE Sensors Journal*, *21*(22), 25532-25546.

Abualigah, L., Elaziz, M. A., Sumari, P., Khasawneh, A. M., Alshinwan, M., Mirjalili, S., ... & Gandomi, A. H. (2022). Black hole algorithm: A comprehensive survey. *Applied Intelligence*, *52*(10), 11892-11915.

Abualigah, L., Habash, M., Hanandeh, E. S., Hussein, A. M., Shinwan, M. Al, Zitar, R. A., & Jia, H. (2023). Improved Reptile Search Algorithm by Salp Swarm Algorithm for Medical Image Segmentation. *Journal of Bionic Engineering*, 1–25.

Abualigah, L., Shehab, M., Alshinwan, M., Mirjalili, S., & Abd Elaziz, M. (2021). Ant lion optimizer: a comprehensive survey of its variants and applications. *Archives of Computational Methods in Engineering*, *28*(3), 1397–1416.

Al-Sakran, A., Qutqut, M. H., Almasalha, F., Hassanein, H. S., & Hijjawi, M. (2018). An overview of the Internet of things closed source operating systems. *2018 14th International Wireless Communications \& Mobile Computing Conference (IWCMC)*, 291–297.

Al Shinwan, M., Abualigah, L., Huy, T.-D., Younes Shdefat, A., Altalhi, M., Kim, C., El-Sappagh, S., Abd Elaziz, M., & Kwak, K. S. (2022). An Efficient 5G Data Plan Approach Based on Partially Distributed Mobility Architecture. *Sensors*, *22*(1), 349.

Al Shinwan, M., Abualigah, L., Le, N. D., Kim, C., & Khasawneh, A. M. (2021). An intelligent long-lived TCP based on real-time traffic regulation. *Multimedia Tools and Applications*, *80*(11), 16763–16780.

Al Shinwan, M., & Chul-Soo, K. (2018). A future mobile packet core network based on ip-in-ip protocol. *International Journal of Computer Networks \& Communications (IJCNC) Vol*, *10*.

Al Shinwan, M., Huy, T.-D., & Chul-Soo, K. (2017). A Flat Mobile Core Network for Evolved Packet Core Based SAE Mobile Networks. *Journal of Computer and Communications*, *5*(5), 62–73.

Alsarhan, T., Harfoushi, O., Shdefat, A. Y., Mostafa, N., Alshinwan, M., & Ali, A. (2023). Improved Graph Convolutional Network with Enriched Graph Topology Representation for Skeleton-Based Action Recognition. *Electronics*, *12*(4), 879.

Alsokkar, A., Law, E., Almajali, D., & Alshinwan, M. (2023). The effect of multimodality on customers' decision-making and experiencing: A comparative study. *International Journal of Data and Network Science*, *7*(1), 1–14.

Baldo, N., Miozzo, M., Requena-Esteso, M., & Nin-Guerrero, J. (2011). An open source product-oriented LTE network simulator based on ns-3. *Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 293–298.

Basta, A., Kellerer, W., Hoffmann, M., Hoffmann, K., & Schmidt, E.-D. (2013). A virtual SDN-enabled LTE EPC architecture: A case study for S-/P-gateways functions. *2013 IEEE SDN for Future Networks and Services (SDN4FNS)*, 1–7.

Bilen, T., Canberk, B., & Chowdhury, K. R. (2017). Handover management in software-defined ultra-dense 5G networks. *IEEE Network*, *31*(4), 49–55.

Bouras, C., & Diles, G. (2015). Resource management in 5G femtocell networks. *2015 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA)*, 353–358.

Bruschi, R., Bolla, R., Davoli, F., Zafeiropoulos, A., & Gouvas, P. (2019). Mobile edge vertical computing over 5G network sliced infrastructures: an insight into integration approaches. *IEEE Communications Magazine*, *57*(7), 78–84.

Carneiro, G., Fortuna, P., & Ricardo, M. (2009). FlowMonitor: a network monitoring framework for the network simulator 3 (NS-3). *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools*, 1–10.

Chan, H. A., Yokota, H., Xie, J., Seite, P., & Liu, D. (2011). Distributed and dynamic mobility management in mobile internet: current approaches and issues. *JCM*, *6*(1), 4–15.

Cominardi, L., Giust, F., Bernardos, C. J., & De La Oliva, A. (2017). Distributed mobility management solutions for next mobile network architectures. *Computer Networks*, *121*, 124–136.

De La Oliva, A., Pérez, X. C., Azcorra, A., Di Giglio, A., Cavaliere, F., Tiegelbekkers, D., Lessmann, J., Haustein, T., Mourad, A., & Iovanna, P. (2015). Xhaul: toward an integrated fronthaul/backhaul architecture in 5G networks. *IEEE Wireless Communications*, *22*(5), 32–40.

Foundation, O. N. (2012). Software-defined networking: the new norm for networks. *ONF White Paper*.

Giust, F., Cominardi, L., & Bernardos, C. J. (2015). Distributed mobility management for future 5G networks: overview and analysis of existing approaches. *IEEE Communications Magazine*, *53*(1), 142–149.

Gupta, A. K., Goel, V., Garg, R. R., Thirupurasundari, D. R., Verma, A., & Sain, M. (2021). A fuzzy based handover decision

scheme for mobile devices using predictive model. *Electronics*, *10*(16), 2016.
Hakeem, S. A. A., Hady, A. A., & Kim, H. (2020). Current and future developments to improve 5G-NewRadio performance in vehicle-to-everything communications. *Telecommunication Systems*, *75*(3), 331–353.
Hu, L., Miao, Y., Yang, J., Ghoneim, A., Hossain, M. S., & Alrashoud, M. (2020). If-rans: intelligent traffic prediction and cognitive caching toward fog-computing-based radio access networks. *IEEE Wireless Communications*, *27*(2), 29–35.
Jeon, S., Figueiredo, S., & Aguiar, R. L. (2013). On the impacts of distributed and Dynamic Mobility Management strategy: A simulation study. *2013 IFIP Wireless Days (WD)*, 1–6.
Khan, S. K., Farasat, M., Naseem, U., & Ali, F. (2020). Performance evaluation of next-generation wireless (5G) UAV relay. *Wireless Personal Communications*, *113*(2), 945–960.
Ki, J.-G., Lee, K.-T., & Kim, D.-H. (2014). Modeling and simulation of partially distributed mobility management scheme. *International Journal of Multimedia and Ubiquitous Engineering*, *9*(8), 125–136.
Kosmopoulos, I., Skondras, E., Michalas, A., Michailidis, E. T., & Vergados, D. D. (2022). Handover Management in 5G Vehicular Networks. *Future Internet*, *14*(3), 87.
Koubias, S. A., & Haralabidis, H. C. (2001). Simulated performance evaluation of a MAC-layer hybrid protocol for multi-channel control networks (MITION). *Telecommunication Systems*, *17*(1), 63–92.
Liu, G., & Jiang, D. (2016). 5G: Vision and requirements for mobile communication system towards year 2020. *Chinese Journal of Engineering*, *2016*(2016), 8.
Makaya, C., & Pierre, S. (2008). An analytical framework for performance evaluation of IPv6-based mobility management protocols. *IEEE Transactions on Wireless Communications*, *7*(3), 972–983.
Masa'deh, R., A. AlMajali, D., AlSokkar, A. A. M., Alshinwan, M., & Shehadeh, M. (2023). Antecedents of Intention to Use E-Auction: An Empirical Study. *Sustainability*, *15*(6), 4871.
Modeas, I., Kaloxylos, A., Merakos, L., & Tsolkas, D. (2021). An adaptive and distributed network selection mechanism for 5G networks. *Computer Networks*, *189*, 107943.
Netmanias Technical Document: EMM Procedure 1. Initial Attach: Part 2. Call Flow of Initial Attach. (2014).
Netmanias Technical Document: EMM Procedure 6. Handover without TAU: Part 2. X2 Handover. (2014).
Nguyen, T.-T., Bonnet, C., & Harri, J. (2016). SDN-based distributed mobility management for 5G networks. *2016 IEEE Wireless Communications and Networking Conference*, 1–7.
Project, 3rd Generation Partnership. (2016). Study on control and user plane separation of EPC nodes. *3GPP TR 23.714 V14.0.0*.
Project, 3rd Generation Partnership. (2020). Architecture enhancements for control and user plane separation of EPC nodes. *3GPP TS 23.214 V16.2.0*.
Project, 3rd Generation Partnership. (2021). General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access. *3GPP TS 23.401 V16.10.0*.
Qureshi, K. N., Abdullah, A. H., Kaiwartya, O., Ullah, F., Iqbal, S., & Altameem, A. (2020). Weighted link quality and forward progress coupled with modified RTS/CTS for beaconless packet forwarding protocol (B-PFP) in VANETs. *Telecommunication Systems*, *75*(2), 145–160.
Qutqut, M. H. (2014). *Mobile Small cells in Cellular Heterogeneous Networks.* Queen's University at Kingston, Ontario, Canada.
Rost, P., Mannweiler, C., Michalopoulos, D. S., Sartori, C., Sciancalepore, V., Sastry, N., Holland, O., Tayade, S., Han, B., Bega, D., & others. (2017). Network slicing to enable scalability and flexibility in 5G mobile networks. *IEEE Communications Magazine*, *55*(5), 72–79.
Said, S. B. H., Sama, M. R., Guillouard, K., Suciu, L., Simon, G., Lagrange, X., & Bonnin, J.-M. (2013). New control plane in 3GPP LTE/EPC architecture for on-demand connectivity service. *2013 IEEE 2nd International Conference on Cloud Networking (CloudNet)*, 205–209.
Shinwan, M. A., & Chul-Soo, K. (2017). Enhanced Mobile Packet Core Network Scheme for Next-Generation Mobile Communication Systems. *International Journal of Electronics Communication and Computer Engineering (IJECCE)*, *8*, 56–61.
Sumathi, D., Prakasam, P., Nandakumar, S., & Balaji, S. (2022). Efficient Seamless Handover Mechanism and Mobility Management for D2D Communication in 5G Cellular Networks. *Wireless Personal Communications*, 1–23.
Sun, W., Wang, L., Liu, J., Kato, N., & Zhang, Y. (2020). Movement aware CoMP handover in heterogeneous ultra-dense networks. *IEEE Transactions on Communications*, *69*(1), 340–352.
Uddin, M. Y., Abdeljaber, H. A., & Ahanger, T. A. (2021). Development of a Hybrid Algorithm for efficient Task Scheduling in Cloud Computing environment using Artificial Intelligence. *International Journal of Computers, Communications and Control*, *16*(5).