

A Flexible Architecture for Broadcast Broadband Convergence in Beyond 5G

Rashmi Yadav*, Rashmi Kamran[†], Pranav Jha[†] and Abhay Karandikar^{†§}

Department of Electrical Engineering, Indian Institute of Technology Kanpur, India*

Department of Electrical Engineering, Indian Institute of Technology Bombay, India[†]

Secretary to the Government of India, Department of Science & Technology, New Delhi, India[§]

Email: rashmiy@iitk.ac.in*, rashmi.kamran@iitb.ac.in[†], pranavjha@ee.iitb.ac.in[†], karandi@ee.iitb.ac.in[†]

Abstract—There has been an exponential increase in the usage of multimedia services in mobile networks in recent years. To address this accelerating data demand, mobile networks are experiencing a subtle transformation in their architecture. One of the changes in this direction is the support of Multicast/Broadcast Service (MBS) in the Third Generation Partnership Project (3GPP) Fifth Generation (5G) network. The MBS has been introduced to enhance resource utilization and user experience in 3GPP 5G networks. However, there are certain limitations in the 3GPP 5G MBS architecture, such as the selection of the delivery method (unicast or broadcast) by the core network (may result in sub-optimal radio resource utilization) and no provision for converging non-3GPP broadcast technologies (like digital terrestrial television) with cellular (3GPP 5G) broadband. In this context, we propose a new architecture for the convergence of cellular broadband and non-3GPP broadcast networks. A novelty of the architecture is that it treats signalling exchange with User Equipment (UE) as data (service) which results in improved scalability of mobile networks. The architecture supports enhanced flexibility in choosing a delivery method (3GPP 5G unicast, 3GPP 5G broadcast, or non-3GPP broadcast) for user data. We evaluate the performance of the proposed architecture using process algebra-based simulations, demonstrating a significant reduction in the number of signalling messages exchanged between the UE and the network for MBS session establishment as compared to the 3GPP 5G network.

Index Terms—Fifth Generation (5G) network, Beyond 5G (B5G), Multicast Broadcast Services (MBS), Digital Terrestrial Television (DTT), Broadcast Broadband Convergence.

I. INTRODUCTION

IN recent years, there has been a remarkable rise in multimedia content utilization over mobile networks. As highlighted in the Ericsson mobility report [1], video content, especially in social media and video-on-demand services, comprises the largest and fastest-growing segment of mobile data traffic globally, which accounts for approximately 70% of traffic share in 2022. Highlights from the same report show an annual growth of about 30% by the end of 2028, further increasing the global mobile data traffic's video share to 80%. Moreover, Qualcomm's broadcast report [2] predicts an enormous rise in live streaming content on social media, with approximately 800 million users expected to participate in daily live streams. These statistics reflect the significant impact and importance of multimedia services in the present-day mobile network.

The deployment of Fifth Generation (5G) mobile networks brings the possibility to increase the usage of Multicast/Broadcast Service (MBS). MBS is a crucial use case to address the increasing data demands within the framework of 5G technology. Apropos to this, Release 17 of the Third Generation Partnership Project (3GPP) 5G standards has introduced the support for MBS to enhance the 3GPP 5G architecture. Nevertheless, there are some architectural limitations of the 3GPP 5G MBS support. These include the selection of delivery methods (unicast or broadcast) by the Core Network (CN) that might result in sub-optimal utilization of resources, limited handling for user mobility for MBS, and no provision for the convergence of Non-3GPP Broadcast Networks (N3BNs) within the 3GPP 5G (such as Digital Terrestrial Television (DTT)).

This section provides the literature survey related to MBS architecture and mechanisms. The work in [3] presents various architectural concepts and mechanisms to optimize network loading and traffic patterns for MBS delivery. The paper [4] presents a comprehensive review on the convergence of broadcast and broadband in the 5G network. The authors in [5] proposed an enhanced Next Generation Radio Access Network (NG-RAN) architecture with architectural and functional enhancements to provide the efficient delivery of terrestrial broadcast services. The work in [6] explores a mixed transmission mode that utilizes shared multicast, broadcast, and unicast resources over the same physical channel. In [7], authors review the upcoming 3GPP 5G standards, discuss limitations of 3GPP 5G MBS architecture and present state-of-the-art standardization initiatives towards integrating N3BNs with the 5G. Furthermore, the latest release of the 3GPP standard (Release 18) does not include the support for integration of the N3BN and the 3GPP broadcast network.

To the best of our knowledge, the prior art lacks comprehensive architectural solutions related to the convergence of cellular broadband (3GPP 5G) and N3BNs and scalability enhancements in the context of MBS delivery in mobile networks. We propose a flexible architecture for broadcast broadband convergence where we treat UE signalling as data or service. Therefore, we call the proposed architecture a Signalling Service-Based Architecture (SSBA) for broadcast broadband convergence. This work is based on prior research conducted in [8], focusing on the convergence of broadcast and

broadband as the authors did not delve into this specific aspect. The SSBA presents a scalable network architecture making it a promising solution for the Beyond 5G (B5G) landscape through enhanced resource utilization in a converged network. SSBA provides the flexibility to choose between broadcast and broadband delivery methods based on resource availability. We are exploring LTE and 5G broadcast (FeMBMS), but the proposed architecture can also incorporate other broadcast schemes, such as the Advanced Television Systems Committee (ATSC). An illustration of the convergence of N3BN in the proposed SSBA is provided in Section V. In Section IV, we evaluate the performance of the proposed SSBA using the Eclipse plug-in [9], a tool for modelling distributed systems with the help of Performance Evaluation Process Algebra (PEPA) [10], a modelling language.

The rest of the paper is as follows: Section II also presents the architectural details of the proposed signalling service-based architecture. Section III presents the system model. We conclude in Section VI along with future directions.

II. PROPOSED SIGNALLING SERVICE-BASED ARCHITECTURE

In this section, we present an overview of the proposed SSBA, as shown in Fig. 1. Before delivering the details of the proposed SSBA, let us first understand the basic principles of an SSBA (a detailed explanation is available in [8]). In SSBA, a Service Function (SF) handles signalling exchange with the UE through the user plane for network services (i.e. RRC/NAS service, broadcast service). The SF then communicates with the network control plane to establish the data path through the network data plane.

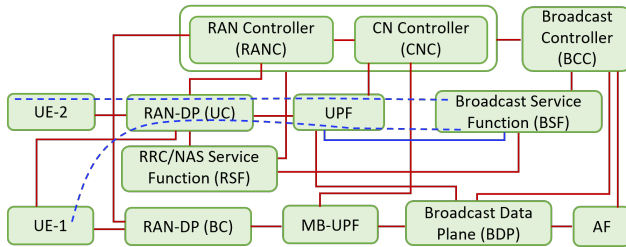


Fig. 1. Proposed signalling service-based mobile network architecture.

In the context of Fig.1, the preceding explanation can be linked as follows: the Broadcast Service Function (BSF) functions as an independent SF responsible for exchanging multicast/broadcast-related signalling messages with UEs. These include membership requests, content delivery requests, and other associated operations. The BSF interfaces with the control plane (Broadcast Controller (BCC)) to initiate data path establishment. Further, the BCC communicates with the RAN/CN controller to configure the data path for efficient content delivery. However, the decision for delivery methods (unicast or broadcast) is governed by BCC and switching between delivery methods is handled by the Broadcast Data Plane (BDP).

The information flow in the proposed SSBA is as follows: UE-1 is associated with the broadcast RAN (RAN-DP (BC)), whereas UE-2 is connected to the unicast (UC) RAN (RAN-DP (UC)). Notably, the BSF facilitates UE (both UE-1 and UE-2) interaction via the DP. Therefore, the signalling path for UC/BC transmission involves the following: UE-2/UE-1 - RAN-DP (UC) - User Plane Function (UPF) - BSF as shown in Fig.1 with the blue dotted line. Furthermore, the RAN controller controls the RAN data planes (RAN-DP (UC) and RAN-DP (BC)), while the CN controller supervises the CN data planes (UPF and Multicast/Broadcast UPF (MB-UPF)).

A. MBS Session Establishment Call flow for the Proposed SSBA

Fig. 2 illustrates the MBS session establishment call flow for the proposed SSBA. Conversely, the call flow for MBS session establishment in the 3GPP 5G architecture is available in [11] (Section 7.1.1.2 and 7.2.1.3). This section provides a concise comparison between the call flows of the proposed and 3GPP 5G MBS session establishment, highlighting the significant reduction in the number of signalling messages achieved in the proposed SSBA.

A Temporary Mobile Group Identity (TMGI) allocation process (messages 1-2) is carried out between Application Function (AF) and BCC as shown in Fig.2. The session create request (message 3) is sent to BCC by AF. In response, BCC forwards this request to the CN controller. CN controller confirms the resource availability from the RAN controller (messages 5 and 7) and the RAN controller sends the session command (message 6) to RAN-DP to set up the session on the RAN side. On receiving confirmation from the RAN controller, the CN controller sends the session command to MB-UPF (message 8) to set up the session on the core network side as well. Messages 13-15 are related to UE joining the MBS session. Here, UE initiates the MBS session join request to BSF, which is then forwarded to the RRC/NAS SF and eventually processed by this SF to send the Radio Resource Control (RRC) reconfiguration message in response, to UE via RAN-DP.

The controllers in the proposed SSBA do not call for response messages from the user plane, as they have global information about the user plane resources. Therefore, messages 6, 8 and 11 (as shown in Fig. 2) are sent as commands without requiring corresponding response messages. As a result, the modified call flow significantly reduces the total number of signalling messages. Similarly, message 14 (Fig. 2) also does not require a response message. Therefore, the procedure for MBS session establishment call flow is appreciably simplified, enhancing the overall modularity of the network.

III. SYSTEM MODEL

In this section, we describe the system modelling of the proposed SSBA, using the PEPA [10], a high-level language, for modelling distributed systems. The PEPA modelling facilitates system performance evaluation using the Eclipse plug-in

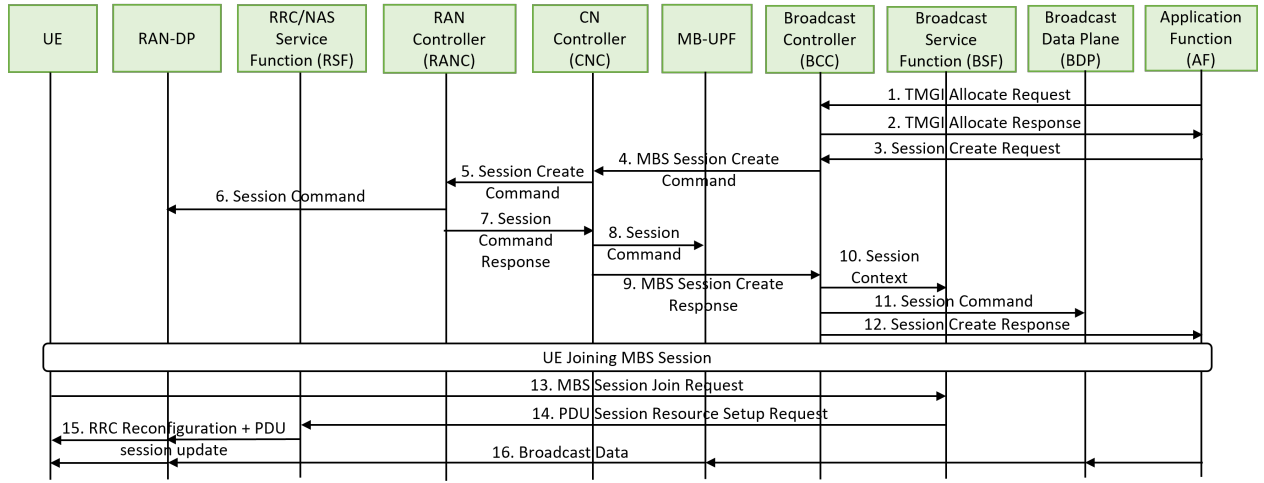


Fig. 2. MBS session establishment call flow in the proposed SSBA.

[9], a tool for performance analysis. In Table I, we present the MBS session establishment call flow (Fig. 2) modelling.

To explain the behaviour of various Network Functions (NFs) involved in the MBS session establishment call flow, we model each NF as a PEPA component. This approach allows us to represent the tasks executed by each PEPA component/NF through different states. Each state is denoted by the NF's name followed by a number corresponding to its state (NF_1). For instance, the NF *UE* possesses two states, Ue_1 and Ue_2 , representing different states of the UE to perform various tasks. As shown in the table, similar representations of states are provided for each NF, such as RAN-DP, SF, RAN controller, CN controller, MB-UPF, BCC, BSF, BDP, and AF.

Furthermore, we model each task of the NF as an action type, indicated in lowercase letters. To provide further details, subscripts are added to the action types ($actiontype_{detail}$). For instance, the action type $tmgi_{req}$ signifies the TMGI allocation request, where $tmgi$ is the action type, and req corresponds to its specific detail as a *request*. Each action type is also associated with a rate value, denoted as r . These rates represent the expected duration of a particular action type in the PEPA component and values are considered as given in [12], [13], and [14].

By using PEPA modelling, we can analyze the performance of the proposed SSBA for MBS session establishment call flow. To understand the system modelling, consider an example of SF as an NF to model it as a PEPA component as shown in Table I. This NF (SF) is represented with two states: Sf_1 and Sf_2 , each performing specific tasks (as shown in Fig. 2). In the first state, Sf_1 , the action type $setup_{req}$ signifies the "PDU session resource setup request" as a task performed during this state.

Moving to the second state, Sf_2 , two actions are being performed within this state. The first action, denoted as get_{sfp} , represents the SF NF's attempt to access the SF processor (SFP) to process the received request. The second action, denoted as $reconfig$, represents the specific action of RRC reconfiguration involving the PDU session modification

command. For the processing of requests, each NF is associated with a corresponding processor. For this purpose, we define processors using a two-state model as defined in [15], [16]. However, the following are the associated processors considered in the example: UE processor (*UEP*), RAN-DP processor (*RANDPP*), *SFP*, RAN controller processor (*RANP*), CN controller processor (*CNP*), MB-UPF processor (*MBUPFP*), BCC processor (*BCCP*), BSF processor (*BSFP*), BDP processor (*BDPP*) and AF processor (*AFP*). For instance, the AFP is defined with two states. The first state, Afp_1 , represents the action of getting access to the AF processor (get_{afp}), while the second state is dedicated to performing actions associated with the processor ($tmgireq$ and $sessionreq$). Similarly, other processors (as shown in Table I) are defined with their corresponding NFs.

The system equation (as shown in Table I) describes the interactions between the NFs in the proposed SSBA. These interactions are performed between various NFs as $NF_1 \bowtie_V NF_2$, where $V = \langle action_1, action_2 \rangle$ is a set of actions performed between NF_1 and NF_2 . For instance, consider the example $Randp_1[N] \bowtie_{V_2} Ran_1[N]$ which signifies that NF Ran_1 interacts with $Randp_1$ NF through the set of actions contained in variable V_2 . Given $V_2 = \langle sessionrandp \rangle$, implies the interaction between Ran_1 and $Randp_1$ to perform the action type $sessionrandp$ (signifies that the RAN controller commands the RAN-DP to set up a session). In a similar way, the interactions between other NFs are modelled and presented in Table I.

Moreover, several other variables are used in the system equation that require further discussion. The variable n represents the number of UEs. N_{nf} denotes the number of specific NFs, such as N_{randp} , N_{sf} , N_{ran} , N_{cn} , N_{mbupf} , N_{bcc} , N_{bsf} , N_{bdp} , and N_{af} , representing the number of RAN-DP, SF, RAN controller, CN controller, MB-UPF, BCC, BSF, BDP, and AF NFs, respectively. Furthermore, each processor can handle a set of concurrent threads, denoted as N_t , while the number of processors allocated to each NF is represented as N_{nfp} . Consequently, $N = N_{nf} \cdot N_{nfp} \cdot N_t$ presents the total number

of threads for a specific NF, and $N_p = N_{nf} \cdot N_{nfp}$ denotes the total number of processors allocated to a particular NF type.

However, we followed a similar modelling procedure and simulations were then performed for both the 3GPP 5G and the proposed SSBA's MBS session establishment call flow, demonstrating a comparative analysis of their performance in terms of the reduced number of signalling messages, enhanced modularity and improved scalability.

IV. PERFORMANCE EVALUATION

In this section, we present the performance evaluation of both the proposed SSBA and the 3GPP 5G architecture using the Eclipse plug-in tool [9]. The evaluation is based on several parameters, such as the number of MBS sessions established per unit time, average response time (ART), and processor utilization. These parameters are significant in evaluating the network's scalability, one of the key aspects we consider. The MBS session establishment rate measures the rate at which MBS sessions are established with respect to specific actions, such as *reconfig* representing RRC reconfiguration (PDU session modification command). This specific action represents the completion of the MBS session establishment call flow. ART evaluates the average waiting time for UE's MBS session establishment process. Processor utilization evaluates the NF's processor capacity utilization during the entire process.

However, with separate controllers, user plane functions, broadcast service functions and other NFs, the proposed SSBA can be viewed as a distributed system similar to the 3GPP 5G architecture, we propose to use the scalability parameter of a distributed system to evaluate and perform the comparison of their scalability. The scalability (S) for a distributed system, based on productivity as defined in [17], is given by the ratio of the productivity of the system at two different configurations with different scales b_1 and b_2 [13]. In this context, the configurations (b_1 and b_2) refer to the different numbers of NFs used in the network. For example, $b_1 = (1,1,1,1,1,1,1,1,1)$ and $b_2 = (3,3,3,3,3,3,3,3,3)$. Configuration b_1 represents the basic configuration with a single NF assigned to RAN-DP, SF, RAN controller, CN controller, BCC, MB-UPF, BSF, BDP and AF in the proposed SSBA (i.e., $N_{randp}, N_{sf}, N_{ran}, N_{cn}, N_{mbupf}, N_{bcc}, N_{bsf}, N_{bdp}, N_{af}$) = $(1,1,1,1,1,1,1,1,1)$. On the other hand, configuration b_2 represents a scaled system with three NFs assigned as follows: $(N_{randp}, N_{sf}, N_{ranc}, N_{cnc}, N_{mbupf}, N_{bcc}, N_{bsf}, N_{bdp}, N_{af}) = (3,3,3,3,3,3,3,3,3)$. Please note that we provide an equal number of total processors in each case (3GPP 5G and the proposed SSBA). The mathematical expression for scalability can be referred from [18]. By evaluating scalability, we can compare the performance of the proposed SSBA and the 3GPP 5G architecture under different scaling configurations which demonstrate the efficient handling of the proposed SSBA for the increased number of users.

Fig. 3 and 4 illustrate the number of MBS sessions established per unit time for both architectures under two different configurations, denoted as b_1 and b_2 . It is observed that the proposed SSBA achieves a higher saturation point than the

TABLE I
SYSTEM MODELLING FOR MBS SESSION ESTABLISHMENT CALL FLOW

PEPA Modelling of NFs	Code Description
UE NF	$Ue_1 \stackrel{def}{=} (get_{uep}, r_p).(mbsjoinreq, r_{iat}).Ue_2$ $Ue_2 \stackrel{def}{=} (reconfig_1, r_v).Ue_1$
RAN-DP NF	$Randp_1 \stackrel{def}{=} (sessionrandp, r_v).Randp_2$ $Randp_2 \stackrel{def}{=} (get_{randpp}, r_p).(prepare, r_v).Randp_1$
RSF NF	$Rsf_1 \stackrel{def}{=} (setupreq, r_v).Rsf_2$ $Rsf_2 \stackrel{def}{=} (get_{rsfp}, r_p).(reconfig, r_v).Rsf_1$
RAN controller NF	$Ran_1 \stackrel{def}{=} (sessioncreate, r_v).Ran_2$ $Ran_2 \stackrel{def}{=} (get_{ranp}, r_p).(sessionrandp, v).(sessionres, v).Ran_1$
CN controller NF	$Cn_1 \stackrel{def}{=} (mbsessioncom, r_v).Cn_2$ $Cn_2 \stackrel{def}{=} (get_{cnp}, r_p).(sessioncreate, r_v).Cn_3$ $Cn_3 \stackrel{def}{=} (sessionres, r_v).Cn_4$ $Cn_4 \stackrel{def}{=} (get_{cnp}, r_p).(sessionupf, r_v).(mbsessionres, r_v).Cn_1$
MB-UPF NF	$Mbupf_1 \stackrel{def}{=} (sessionupf, r_v).Mbupf_2$ $Mbupf_2 \stackrel{def}{=} (get_{mbupfp}, r_p).(prepare, r_v).Mbupf_1$
BCC NF	$Bcc_1 \stackrel{def}{=} (tmgireq, r_v).Bcc_2$ $Bcc_2 \stackrel{def}{=} (get_{bccp}, r_p).(tmgires, r_v).Bcc_3$ $Bcc_3 \stackrel{def}{=} (sessionreq, r_v).Bcc_4$ $Bcc_4 \stackrel{def}{=} (get_{bccp}, r_p).(mbsessioncomm, r_v).$ $Bcc_4 \stackrel{def}{=} (mbsessionres, r_v).Bcc_5$ $Bcc_5 \stackrel{def}{=} (get_{bccp}, r_p).(sessioncontext, r_v).(sessionbdp, r_v).(sessionres, r_v).Bcc_1$
BSF NF	$Bsf_1 \stackrel{def}{=} (sessioncontext, r_v).(mbsjoinreq, r_v).Bsf_2$ $Bsf_2 \stackrel{def}{=} (get_{bsfp}, r_p).(setupreq, r_v).Bsf_1$
BDP NF	$Bdp_1 \stackrel{def}{=} (sessionbdp, r_v).Bdp_2$ $Bdp_2 \stackrel{def}{=} (get_{bdpp}, r_p).(prepare, r_v).Bdp_1$
AF NF	$Af_1 \stackrel{def}{=} (get_{afp}, r_p).(tmgireq, r_v).Af_2$ $Af_2 \stackrel{def}{=} (tmgires, r_v).Af_3$ $Af_3 \stackrel{def}{=} (get_{afp}, r_p).(sessionreq, r_v).Af_4$ $Af_4 \stackrel{def}{=} (sessionres, r_v).Af_1$
UE Processor	$Uep_1 \stackrel{def}{=} (get_{uep}, r_p).Uep_2$ $Uep_2 \stackrel{def}{=} (mbsjoinreq, r_{iat}).Uep_1$
RAN-DP Processor	$Randpp_1 \stackrel{def}{=} (get_{randpp}, r_p).Randpp_2$ $Randpp_2 \stackrel{def}{=} (prepare, r_v).Randpp_1$
RSF Processor	$Rsf_1 \stackrel{def}{=} (get_{rsfp}, r_p).Rsf_2$ $Rsf_2 \stackrel{def}{=} (reconfig, r_v).Rsf_1$
RAN Processor	$Ranp_1 \stackrel{def}{=} (get_{ranp}, r_p).Ranp_2$ $Ranp_2 \stackrel{def}{=} (sessionrandp, r_v).Ranp_1 + (sessionres, r_v).Ranp_1$
CN Processor	$Cnp_1 \stackrel{def}{=} (get_{cnp}, r_p).Cnp_2$ $Cnp_2 \stackrel{def}{=} (sessioncreate, r_v).Cnp_1 + (sessionupf, r_v).Cnp_1 + (mbsessionres, r_v).Cnp_1$
MB-UPF Processor	$Mbupfp_1 \stackrel{def}{=} (get_{mbupfp}, r_p).Mbupfp_2$ $Mbupfp_2 \stackrel{def}{=} (prepare, r_v).Mbupfp_1$
BCC Processor	$Bccp_1 \stackrel{def}{=} (get_{bccp}, r_p).Bccp_2$ $Bccp_2 \stackrel{def}{=} (tmgires, r_v).Bccp_1 + (mbsessioncom, r_v).Bccp_1 + (sessioncontext, r_v).Bccp_1 + (sessionres, r_v).Bccp_1 + (sessionbdp, r_v).Bccp_1$
BSF Processor	$Bsf_1 \stackrel{def}{=} (get_{bsfp}, r_p).Bsf_2$ $Bsf_2 \stackrel{def}{=} (setupreq, r_v).Bsf_1$
BDP Processor	$Bdpp_1 \stackrel{def}{=} (get_{bdpp}, r_p).Bdpp_2$ $Bdpp_2 \stackrel{def}{=} (prepare, r_v).Bdpp_1$
AF Processor	$Afp_1 \stackrel{def}{=} (get_{afp}, r_p).Afp_2$ $Afp_2 \stackrel{def}{=} (tmgireq, r_v).Afp_1 + (sessionreq, r_v).Afp_1$
System Equation	$Ue_1[n]_{\phi}^{\boxtimes} Randp_1[N]_{V_1}^{\boxtimes} Rsf_1[N]_{V_2}^{\boxtimes} Ran_1[N]_{V_3}^{\boxtimes} Cn_1[N]_{V_4}^{\boxtimes} Mbupf_1[N]_{V_5}^{\boxtimes} Bcc_1[N]_{V_6}^{\boxtimes} Bsf_2[N]_{V_7}^{\boxtimes} Bdp_1[N]_{V_8}^{\boxtimes} Af_2[N]_{V_9}^{\boxtimes} Uep_1[n]_{\phi}^{\boxtimes} Randpp_1[N_p]_{\phi}^{\boxtimes} Rsf_1[N_p]_{\phi}^{\boxtimes} Ranp_1[N_p]_{\phi}^{\boxtimes} Cnp_2[N_p]_{\phi}^{\boxtimes} Mbupfp_1[N_p]_{\phi}^{\boxtimes} Bccp_1[N_p]_{\phi}^{\boxtimes} Bsf_2[N_p]_{\phi}^{\boxtimes} Bdpp_1[N_p]_{\phi}^{\boxtimes} Af_2[N_p]_{\phi}^{\boxtimes}$
Variables	$V_1 = \langle reconfig \rangle$ $V_2 = \langle sessionrandp \rangle$ $V_3 = \langle sessioncreate, sessionres \rangle$ $V_4 = \langle sessionupf \rangle$ $V_5 = \langle mbsessioncom, mbsessionres \rangle$ $V_6 = \langle mbsjoinreq, setupreq \rangle$ $V_7 = \langle sessionbdp \rangle$ $V_8 = \langle tmgireq, tmgires, sessionreq, sessionres \rangle$ $V_9 = \langle get_{uep}, get_{randpp}, get_{rsfp}, get_{ranp}, get_{cnp}, get_{mbupfp}, get_{bccp}, get_{bsfp}, get_{bdpp}, get_{afp} \rangle$ $\phi = \langle \rangle$

3GPP 5G architecture. In the basic configuration (b_1), the 3GPP 5G architecture saturates at 14,000 users, while the proposed SSBA saturates at 30,000 users. Similarly, in the scaled configuration (b_2), the 3GPP 5G architecture saturates at 42,000 users, while the proposed SSBA saturates at 90,000 users. The saturation point indicates the maximum number of UEs served by the network before it becomes overloaded.

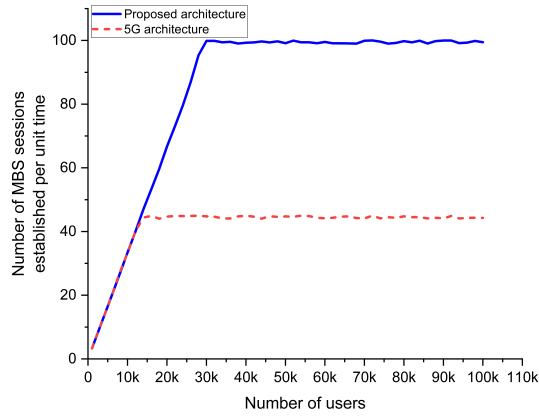


Fig. 3. Number of MBS sessions established per unit time for the basic configuration.

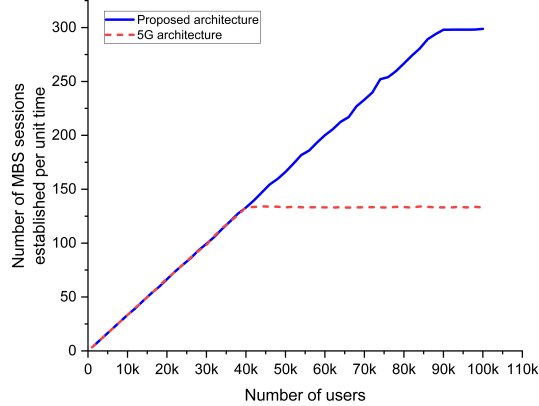


Fig. 4. Number of MBS sessions established per unit time for the scaled configuration.

The comparative analysis of processor utilization of both the 3GPP 5G and the proposed SSBA for the basic configuration is shown in Fig. 5. It shows that the BCCP (BCC Processor), for instance, reaches its maximum processor utilization, explaining the saturation point for the number of MBS session establishments. Other NFs, however, are not fully utilized at this point. This signifies that the processing chain fails if an NF becomes a bottleneck in the consecutive chain. Fig. 6 presents the processor utilization results for both architectures for the scaled configuration. The results demonstrate that processors in the 3GPP 5G architecture saturate earlier than the proposed SSBA due to the higher number of messages in the 3GPP 5G architecture.

Based on the obtained results for the MBS session rate, ART, and processor utilization, the scalability is evaluated

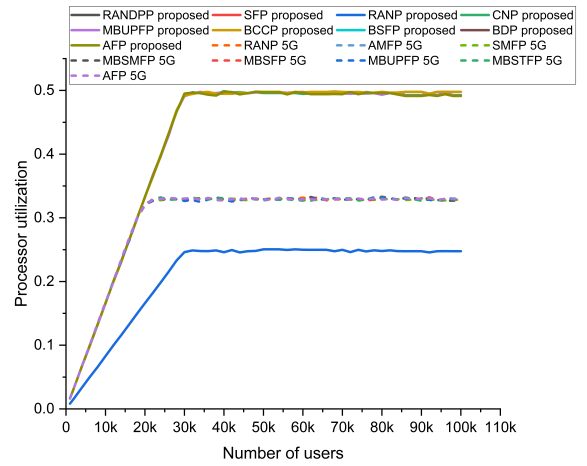


Fig. 5. Processor utilization of MBS session establishment for the basic configuration.

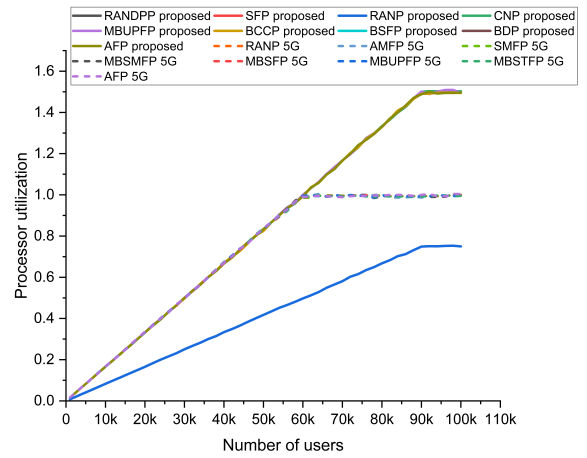


Fig. 6. Processor utilization of MBS session establishment for the scaled configuration.

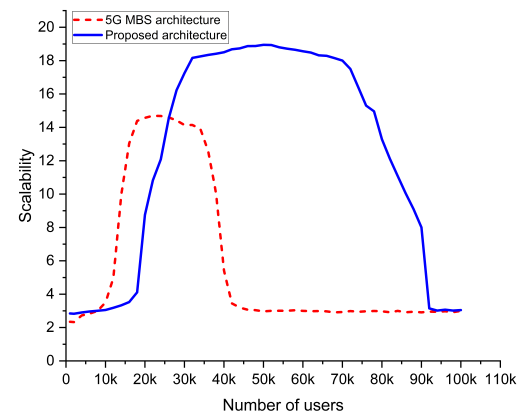


Fig. 7. MBS session establishment scalability.

using the equation provided in [18]. The scalability results are plotted in Fig. 7 for configurations b_1 and b_2 . The proposed SSBA outperforms the 3GPP 5G architecture, as it can serve more concurrent users with the same scaling configuration. It

is evident from the results that the proposed SSBA is more scalable and performs better than the 3GPP 5G architecture.

V. CONVERGENCE OF 3GPP 5G AND N3BN IN SSBA

The proposed SSBA can also be extended to converge 3GPP 5G cellular broadband and N3BN (say, DTT) easily. The simplified integration of N3BN (DTT) with 3GPP 5G is shown in Fig. 8. The MBS-related signalling is taken care of by the BSF. However, the data path in the case of DTT content delivery is as follows: AF - BDP - DTT CN - DTT RAN - UE-3. The BSF serves as the integration point for the signalling interplay between 3GPP 5G and the DTT broadcasting network, while the BDP emerges as the point for data path integration.

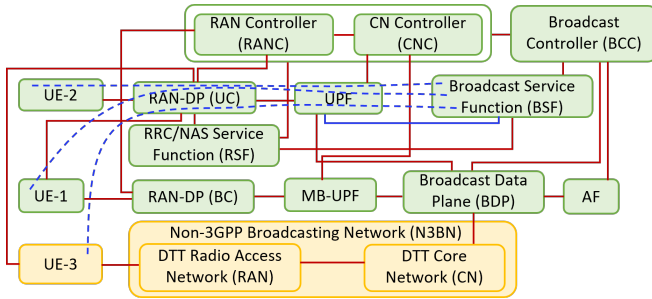


Fig. 8. Convergence of N3BN (DTT) in the proposed SSBA.

VI. CONCLUSION

In this paper, we have proposed a signalling service-based architecture for MBS that offers enhanced flexibility to select a delivery method based on resource availability, and results in improved network scalability by handling UE signalling as a service. Besides, it also facilitates the convergence of 3GPP 5G cellular broadband and N3BNs in the landscape of 5G. The simulations and performance evaluations were performed to demonstrate that the proposed SSBA outperforms 3GPP 5G architecture, exhibiting enhanced modularity, scalability, and a reduced number of signalling messages in the MBS session establishment procedure. In the future, we would like to perform the evaluation of the converged SSBA mobile network.

ACKNOWLEDGMENT

We acknowledge the Ministry of Electronics and Information Technology (MeitY), Govt. of India for supporting the project.

REFERENCES

- [1] Ericsson, "Ericsson Mobility Report," *Report*, November 2022.
- [2] Qualcomm, "Pioneering 5G Broadcast - Building on multiple generations of cellular broadcast technology leadership," *Report*, May 2021.
- [3] A. Saha, M. Kulkarni, P. Naik, A. Chakraborty, G. Padaki and S. KS, "A Collaborative RAN Approach for Handling Multicast-Broadcast Traffic in 5GS," *IEEE 3rd 5G World Forum (5GWF)*, pp. 92–97, 2020.
- [4] Gomez-Barquero, David and Lee, Jae-Young and Ahn, Sungjun and Akamine, Cristiano and He, Dazhi and Montalaban, Jon and Wang, Jintao and Li, Wei and Wu, Yiyang, "IEEE Transactions on Broadcasting: Convergence of Broadcast and Broadband in the 5G Era," *IEEE Transactions on Broadcasting*, vol. 66, no. 2, pp. 383–389, 2020.

- [5] Säily, Mikko and Estevan, Carlos Barjau and Gimenez, Jordi Joan and Tesema, Fasil and Guo, Wei and Gomez-Barquero, David and Mi, De, "5G Radio Access Network Architecture for Terrestrial Broadcast Services," *IEEE Transactions on Broadcasting*, vol. 66, no. 2, pp. 404–415, 2020.
- [6] Guo, Wei and Fuentes, Manuel and Christodoulou, Louis and Mouhouche, Belkacem, "Roads to Multimedia Broadcast Multicast Services in 5G New Radio," in *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2018, pp. 1–5.
- [7] Kamran, Rashmi and Jha, Pranav and Kiran, Shwetha and Karandikar, Abhay and Chaporkar, Prasanna and Saha, Anindya and Chakraborty, Arindam, "A Survey on Multicast Broadcast Services in 5G and Beyond," in *2022 National Conference on Communications (NCC)*, 2022, pp. 344–349.
- [8] Khaturia, Meghna and Manjeshwar, Akshatha Nayak and Jha, Pranav and Karandikar, Abhay, "5G-Serv: Decoupling User Control and Network Control in the 3GPP 5G Network," in *24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, 2021, pp. 75–79.
- [9] M. Tribastone, "The PEPA Eclipse Plugin," *SIGMETRICS Perform. Eval. Rev.*, vol. 36, no. 4, p. 28–33, 2009.
- [10] Hillston, Jane, "A Compositional Approach to Performance Modelling," *Thesis, Cambridge University Press*, 1996.
- [11] 3GPP TS 23.247, "Architectural enhancements for 5G multicast-broadcast services; Stage 2 (V17.1.0 (2021-12))," *Technical Specification*, 2021.
- [12] Oliveira, E. M. R and Viana, A. C. and Naveen, K. P. and Sarraute, C., "Mobile data traffic modeling: Revealing temporal facets," *Computer Networks*, vol. 112, p. 176–193, 2017.
- [13] Arteaga, Carlos Hernan Tobar and Ordoñez, Armando and Rendon, Oscar Mauricio Caicedo, "Scalability and Performance Analysis in 5G Core Network Slicing," *IEEE Access*, vol. 8, pp. 142086–142100, 2020.
- [14] Arteaga, Carlos Hernan Tobar, "A 5G Core Network Prototype," *Computer Networks*, vol. 112, p. 176–193, 2019.
- [15] Christopher D. Williams and Allan Clark, "A Case Study in Capacity Planning for PEPA Models with the PEPA Eclipse Plug-in," *Electronic Notes in Theoretical Computer Science*, vol. 318, p. 69–89, 2015.
- [16] Hillston, Jane and Tribastone, Mirco and Gilmore, Stephen, "Stochastic Process Algebras: From Individuals to Populations," *The Computer Journal*, vol. 55, no. 7, p. 866–881, 2012.
- [17] Jogalekar, Prasad and Woodside, Murray, "Evaluating the scalability of distributed systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 11, no. 6, p. 589–603, 2000.
- [18] Yadav, Rashmi and Kamran, Rashmi and Jha, Pranav and Karandikar, Abhay, "An Architecture for Control Plane Slicing in Beyond 5G Networks," 2023. [Online]. Available: 10.48550/arXiv.2307.06262.