DOI: 10.3969/j.issn.1673-5188.2018.02.005
http://kns.cnki.net/kcms/detail/34.1294.TN.20180606.1754.002.html, published online June 6, 2018

General Architecture of Centralized Unit and Distributed Unit for New Radio

GAO Yin, HAN Jiren, LIU Zhuang, LIU Yang, and HUANG He
(Algorithm Department, ZTE Corporation, Shanghai 201203, China)

Abstract

In new radio (NR) access technology, the radio access network (RAN) architecture is split into two kinds of entities, i.e., the centralized unit (CU) and the distributed unit (DU), to enhance the network flexibility. In this split architecture, one CU is able to control several DUs, which enables the function of base-band central control and remote service for users. In this paper, the general aspects of CU-DU split architecture are introduced, including the split method, interface functions (control plane functions and user plane functions), mobility scenarios and other CU-DU related issues. The simulations show the performance of Options 2 and 3 for CU-DU split.

Keywords

NR; CU; DU; F1 interface

1 Introduction

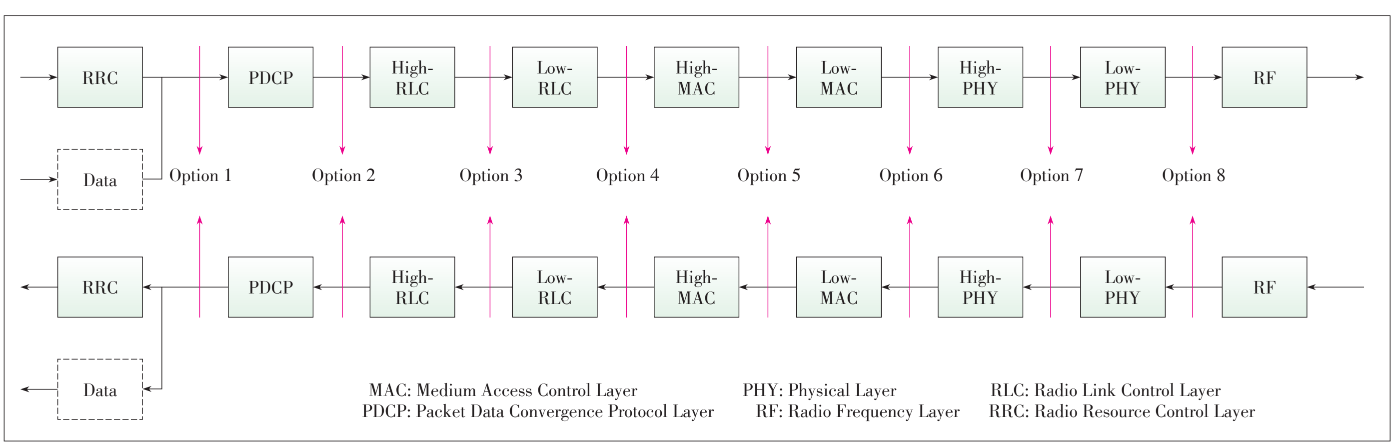
There are transport networks with performance that varies from high transport latency to low transport latency in real deployment. In order to cater for these various types of transport networks and realize multi-vendor CU-DU operation, the radio access network (RAN) architecture for new radio (NR) is split into two kinds of entities, i.e., the centralized unit (CU) and the distributed unit (DU). The latency-tolerant network function resides in the CU entity, and the latency-sensitive network function resides in the DU entity [1].

Fig. 1 shows the possible CU-DU split options [2]. Options 1, 2, 3, 4, and 5 are regarded as higher layer split variants, while Options 6, 7, and 8 are regarded as lower layer split variants in the case of CU-DU.

2 High Layer Split (HLS)

For a transport network with higher transport latency, higher layer splits may be applicable. On the other hand, for a transport network with lower transport latency, lower layer splits can also be applicable.

The choice of how to split functions in the 5G RAN architecture should offer good performance of services. The 3rd Generation Partnership Project (3GPP) agrees that there shall be nor-



▲**Figure 1. Function split between centralized and distributed units [2].**

mative work for a single HLS option (Option 2 or Option 3), and finally Option 2 for high layer RAN architecture split is selected because better performance with more high throughput and less latency restriction can be provided by Option 2 compared with Option 3. The detailed comparison of the two options based on simulation is shown in Section 7.

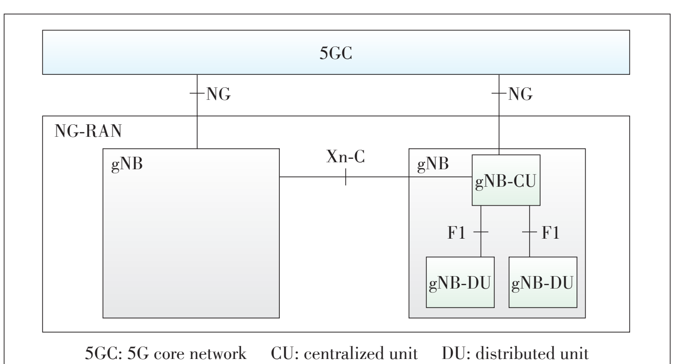
3 Overall Architecture

As shown in **Fig. 2**, in the next generation radio access network (NG-RAN), there are a set of next generation NodeBs (gNBs) connected to the 5G core network (5GC) through the NG interface, and the gNBs can be interconnected through Xn interface. For disaggregate cases, a gNB may consist of a gNB-CU and one or more gNB-DU(s), and the interface between gNB-CU and gNB-DU is called F1. The NG and Xn-C interfaces for a gNB terminate in the gNB-CU. One gNB-CU can connect to multiple gNB-DUs, and the maximum number of connected gNB-DUs is only limited by implementation.

In the 3GPP standard, one gNB-DU is supported to connect only one gNB-CU. However, a gNB-DU can be connected to multiple gNB-CUs by appropriate implementation for resiliency. Meanwhile, one gNB-DU can support one or more cells. The internal structure of the gNB is not visible to the core network and other RAN nodes, the gNB-CU and connected gNB-DUs are only visible to other gNBs and the 5GC as a gNB. With the analysis above, the following definitions of gNB-CU and gNB-DU can be obtained.

The gNB-CU is a logical node hosting Radio Resource Control Layer (RRC), Service Data Adaptation Protocol (SDAP) and PDCP protocols of the gNB or RRC and PDCP protocols of the evolved universal terrestrial radio access-new radio gNB (en-gNB) that controls the operation of one or more gNB-DUs. The gNB-CU terminates the F1 interface connected with the gNB-DU [3].

The gNB-DU is a logical node hosting Radio Link Control Layer (RLC), MAC and PHY layers of the gNB or en-gNB, and its operation is partly controlled by gNB-CU. One gNB-DU supports one or multiple cells. One cell is supported by only one gNB-DU. The gNB-DU terminates the F1 interface connected with the gNB-CU [3].



▲**Figure 2. Overall architecture of NG-RAN [3].**

4 F1 Interface Principle

The interface between gNB-CU and gNB-DU is called F1, and similar to NG or Xn interface in 5G RAN, it supports signalling exchange and data transmission between endpoints. Besides, F1 interface separates the radio network layer and the transport network layer, and it enables exchange of UE associated signalling and non-UE associated signalling. In addition, F1 interface supports control plane (CP) and user plane (UP) separation, therefore, the F1 interface functions are divided into F1-C function and F1-U function.

4.1 F1-C Function

Considering the control plane function of F1 interface, the F1 interface management, system information management, UE context management and RRC message transfer should be introduced.

The F1 interface management function mainly consists of F1 setup, gNB-CU configuration update, gNB-DU configuration update, error indication, and reset function.

The F1 setup function is responsible for the exchange of application level data between gNB-DU and gNB-CU, and it can activate the cells in gNB-DU. The F1 setup procedure is initiated by the gNB-DU. The gNB-CU configuration update and gNB-DU configuration functions are responsible for the update of application level data configuration between gNB-DU and gNB-CU. The gNB-DU configuration update can also activate or deactivate the cells in gNB-DU. Besides, the F1 setup and gNB-DU configuration update functions allow to inform the S-NSSAI(s) supported by the gNB-DU. In addition, the error indication function is responsible for indicating that an error has occurred and reset function is responsible for initializing the peer entity after node setup and after a failure event occurs.

As for system information management, the gNB-DU is responsible for system broadcast information scheduling and system information transmission. For the system information broadcast, the encoding of NR-Master Information Block (MIB) and System Information Block 1 (SIB1) is carried out by the gNB-DU, while the encoding of other system information (SI) messages is carried out by the gNB-CU.

For the sake of UE energy saving, the on-demand SI delivery is introduced over F1 interface as well. In this case, CU is in charge of processing the on-demand SI request from UE over MSG3 and sends System Information Delivery Command to tell gNB-DU broadcast the requested other SI(s), the UE is able to obtain the requested SI(s) from the gNB-DU when needed instead of monitoring the broadcast channel all the time.

The F1 UE context management function is responsible for the establishment and modification of the necessary overall UE context.

The establishment of F1 UE context is initiated by gNB-CU, and gNB-DU can accept or reject the establishment based on admission control criteria (e.g., the resource is not available). Besides, the modification of F1 UE context can be initiated by either gNB-CU or gNB-DU. The receiving node can accept or reject the modification. Furthermore, the F1 UE context management function also supports the release of the context previously established in the gNB-DU. The release of the context is triggered by the gNB-CU either directly or following a request received from the gNB-DU. The gNB-CU requests the gNB-DU to release the UE context when the UE enters RRC_IDLE or RRC_INACTIVE.

The F1 UE context management function can also be used to manage data radio bearers (DRBs) and signaling radio bearers (SRBs), i.e., establishing, modifying and releasing DRB and SRB resources. The establishment and modification of DRB resources are triggered by the gNB-CU and accepted/rejected by the gNB-DU based on resource reservation information and QoS information to be provided to the gNB-DU. For each DRB to be setup or modified, the signal network slice selection assistance information (S-NSSAI) may be provided by gNB-CU to the gNB-DU in the UE context setup procedure and the UE context modification procedure.

The mapping between QoS flows and radio bearers is performed by gNB-CU and the granularity of bearer related management over F1 is radio bearer level. To support PDCP duplication for intra-DU carrier aggregation (CA), one data radio bearer should be configured with two GTP-U tunnels between gNB-CU and gNB-DU [4].

The RRC message transfer function is responsible for the transfer of RRC messages between gNB-CU and gNB-DU. RRC messages are transferred over F1-C, while the UE related RRC messages are transferred over the Uu interface.

4.2 F1-U Function

Considering the user plane function of F1 interface, the user data transfer and flow control should be introduced.

The user data transfer function allows the transferring of user data between gNB-CU and gNB-DU.

The flow control function allows controlling downlink user data transmission towards the gNB-DU. The function includes the transmitting procedure of DL USER DATA and DL DATAT DELIVEREY STATUS frames. There are several methods for the flow control enhancement of data transmission introduced in 3GPP standard [5].

The transfer procedure of downlink user data (DL USER DATA frame) aims to provide F1-U specific sequence number information when transferring user data carrying a DL PDCP PDU from gNB-CU to gNB-DU via the F1-U interface. For the DL USER DATA frame, in order to discard the redundant PDUs caused by the PDCP duplication, the discarded flag and the information on discarding the PDCP PDUs between a start and a stop range are added in the DL USER DATA frame, i.e., the DL discards the NR PDCP PDU SN start (first/last block) and the corresponding discarded block size (first/last block). For retransmitted data packets, a "retransmission flag" is introduced in the spare bit of DL_USER_DATA, which helps the gNB-DU to identify and handle the retransmitted packets with high priority. The gNB-CU can set the Report Polling Flag within the DL USER DATA frame to confirm DL DATA DELIVERY STATUS from the gNB-DU.

After receiving a DL USER DATA frame from the gNB-CU, the gNB-DU shall detect whether an F1-U packet is lost over the F1 interface and memorize the respective sequence number after it declares the respective F1-U packet as being "lost", and the gNB-DU shall transfer the remaining NR PDCP PDUs towards the UE and memorize the highest NR PDCP PDU sequence number of the NR PDCP PDU that has successfully been delivered in sequence towards the UE (in case RLC AM is used) and the highest NR PDCP PDU sequence number of the NR PDCP PDU that has been transmitted to the lower layers. The gNB-DU shall send the DL DATA DELIVERY STATUS if the Report Polling Flag is set.

The transfer procedure of Downlink Data Delivery Status (DL DATA DELIVEREY STATUS frame/DDDS frame) aims to provide feedback from gNB-DU to gNB-CU to allow the gNB-CU to control the downlink user data flow via the gNB-DU for the respective data radio bearer. For the DL DATA DELIVEREY STATUS frame, the highest successfully delivered/transmitted NR PDCP sequence number is added, which can help gNB-CU acquire more accurate data delivery status in the gNB-DU for RLC AM/UM mode data. For the fast data retransmission of lost PDCP PDUs caused by radio link outage, the DL DATA DELIVERY STATUS frame includes the indication of detected radio link outage/resume, together with the information on the highest NR PDCP PDU sequence number successfully delivered in sequence to the UE and the highest NR PDCP PDU sequence number transmitted to the lower layers. The gNB-DU shall indicated lost NR-U packets over the F1 interface within the DDDS frame and also set the desired buffer size for the concerned data bearer and the minimum desired buffer size for the UE within the DDDS frame.

After receiving the DL DATA DELIVERY STATUS frame from the gNB-DU, the gNB-CU shall regard the desired buffer size and the minimum desired buffer size as the amount of data desired from the gNB-DU, and remove the buffered PDCP PDUs according to the feedback of successfully delivered PDCP PDUs. The gNB-CU should also decide the actions necessary for undelivered/transmitted PDCP PDUs at the gNB-DU side, e.g., retransmitting corresponding PDUs to other available gNB-DUs when outage reported.

5 Mobility Scenarios

5.1 Intra-gNB-CU Mobility

In the Intra-gNB-CU mobility part, the standalone case and dual connectivity case are considered.

5.1.1 Intra-NR Mobility

In this scenario, the source and target cells belong to different gNB-DUs in the same gNB-CU.

The gNB-CU makes a decision on the suitable target gNB-DU for handover, based on the UE measurement report. Then, the gNB-CU initiates the UE Context Setup procedure to assign resources on Uu and F1 for one or several RBs and to setup corresponding context for a given UE in the target gNB-DU. The target gNB-DU shall execute the requested RB configuration, and if available, stores the general UE context. At the next step, the gNB-CU sends the RRC reconfiguration message including Cell Group Config at least in the target gNB-DU to the UE. Finally, the UE sets up the RRC connection with the target gNB-DU and replies the RRC reconfiguration complete message. After the UE access to the target gNB-DU, the gNB-CU initiates the UE Context Release procedure to release the UE context in the source gNB-DU [6]. The signaling flow is shown in **Fig. 3** [3].

5.1.2 Inter-gNB-DU Mobility with Dual Connectivity

In this scenario, the source cell and the target cell belong to different gNB-DUs in the same gNB-CU, and a UE can connect with one or more gNB-DUs at the same time.

The UE moves between the cells belonging to different gNB-DUs. The gNB-CU makes a decision on suitable target gNB-DU addition based on the UE measurement report. Then, the gNB-CU initiates the UE Context Setup procedure to assign resources on Uu and F1 for one or several RBs and to setup corresponding context for a given UE in the secondary gNB-DU. The gNB-DU shall execute the requested RB configuration, and if available, stores the general UE context. At the next step, the gNB-CU sends the RRC reconfiguration message including Cell Group Config at least in the secondary gNB-DU to the UE. Finally, the UE sets up the RRC connection with the target gNB-DU (gNB-DU2) and replies the RRC reconfiguration complete message. After the UE context has established in the target gNB-DU (gNB-DU2), UE is connecting with both target gNB-DU (gNB-DU2) and source gNB-DU (gNB-DU1) at the same time. If one leg breaks during the dual connectivity, the fast centralized retransmission procedure of lost PDUs should be used [7]. This signaling flow is shown in **Fig. 4** [8].

5.1.3 Evovled Universal Terrestrial Radio Access-New Radio Dual Connectivity (EN-DC) Mobility

In this scenario, the source cell and the target cell belong to different gNB-DUs in the secondary node.

The Master eNB (MeNB) makes a decision on the suitable target gNB-DU for handover, based on the UE measurement report. Then, after receiving the SgNB Modification Request message from the MeNB with SCG configuration, the gNB-CU initiates the UE Context Setup procedure to assign resources on Uu and F1 for one or several RBs and to setup corresponding context for a given UE in the target gNB-DU. The target gNB-DU shall execute the requested RB configuration, and if available,

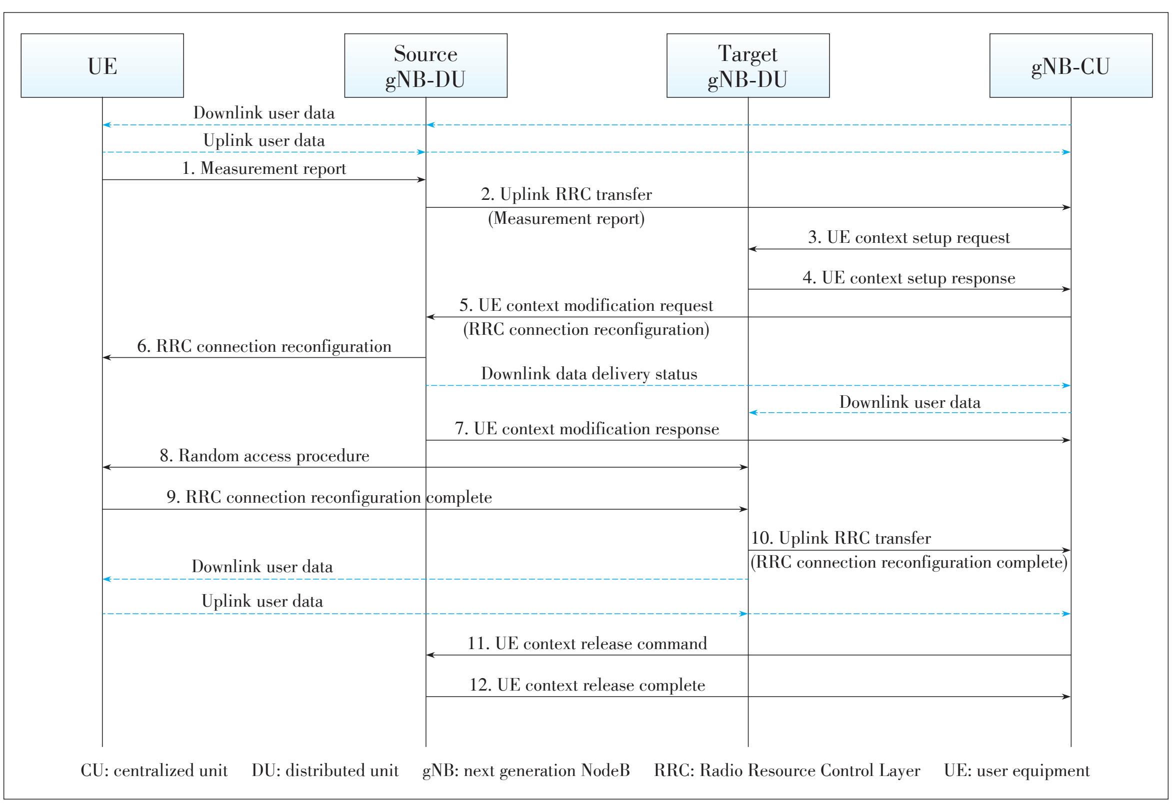
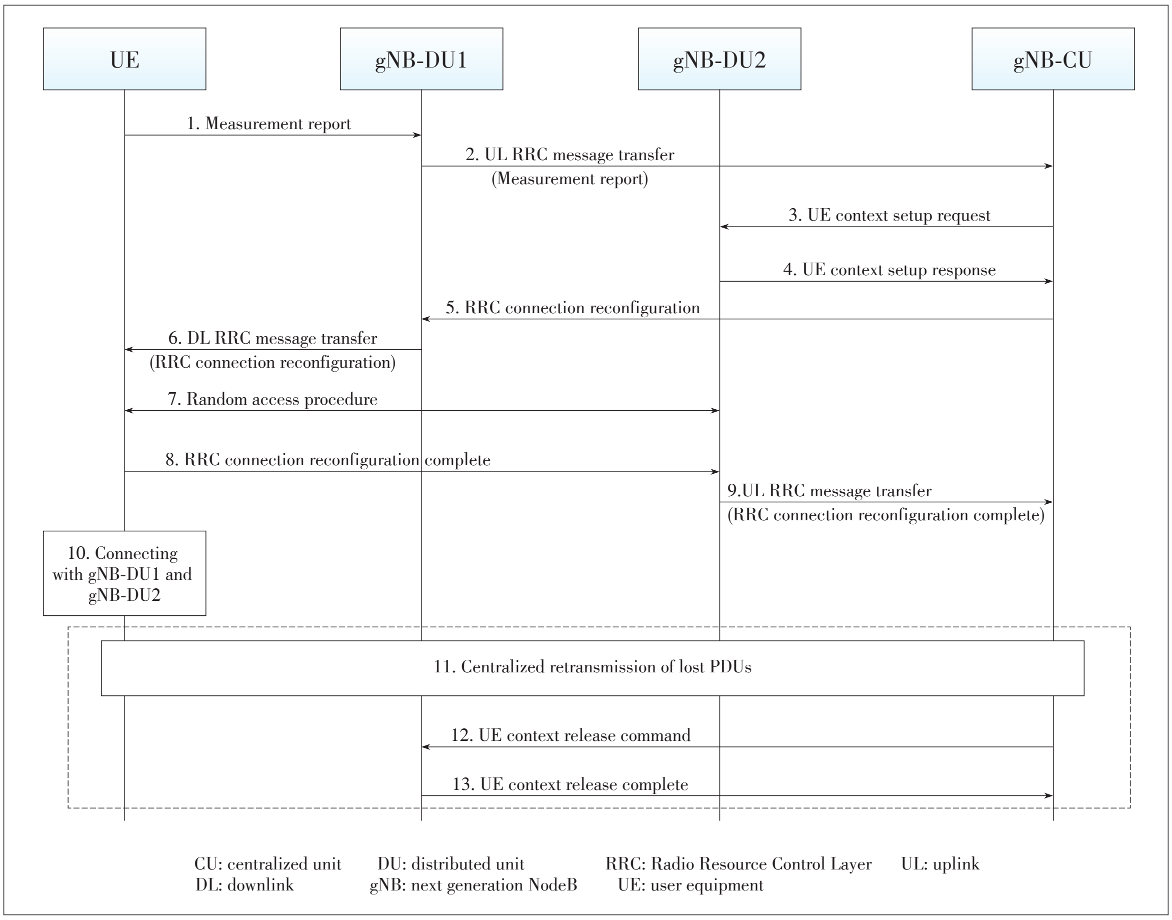


Figure 3. Inter-gNB-DU mobility for intra-NR [3].



◀**Figure 4. Inter-DU handover procedure with dual connectivity [8].**

store the general UE Context. After that, gNB-CU sends the confirmed SCG configuration to the MeNB which needs to be transferred to UE. Finally, the UE sets up the RRC connection with the target gNB-DU. After the UE context has been established in the target gNB-DU, the gNB-CU initiates the UE Context Release procedure to release the UE context in the source gNB-DU [9]. **Fig. 5** shows the signaling flow [3].

6 Other CU-DU Related Issues

6.1 CU-DU Low Layer Split (LLS)

In addition to CU-DU HLS, lower layer split is also applicable and preferable to realize enhanced performance (e.g. centralized scheduling) for transport network with lower transport latency. In this case, the physical layer is split into LLS-CU and LLS-DU. The possible LLS options to be discussed are shown in **Fig. 6** [10].

The possible non-exhaustive functional split options (Fig. 6) for DL and UL are listed as below:

1) Option 6

All of the PHY functions reside in the DU.

2) Option 7-1

In the UL, FFT and CP removal functions reside in the LLS-DU, while the rest of PHY functions reside in the LLS-CU.

In the DL, iFFT and CP addition functions reside in the LLS-DU, while the rest of PHY functions also reside in the LLS-CU.

3) Option 7-2

In the UL, FFT and CP removal and resource de-mapping functions reside in the LLS-DU, while the rest of PHY functions reside in the LLS-CU.

In the DL, iFFT and CP addition, resource mapping and precoding functions reside in the LLS-DU, while the rest of PHY functions reside in the LLS-CU.

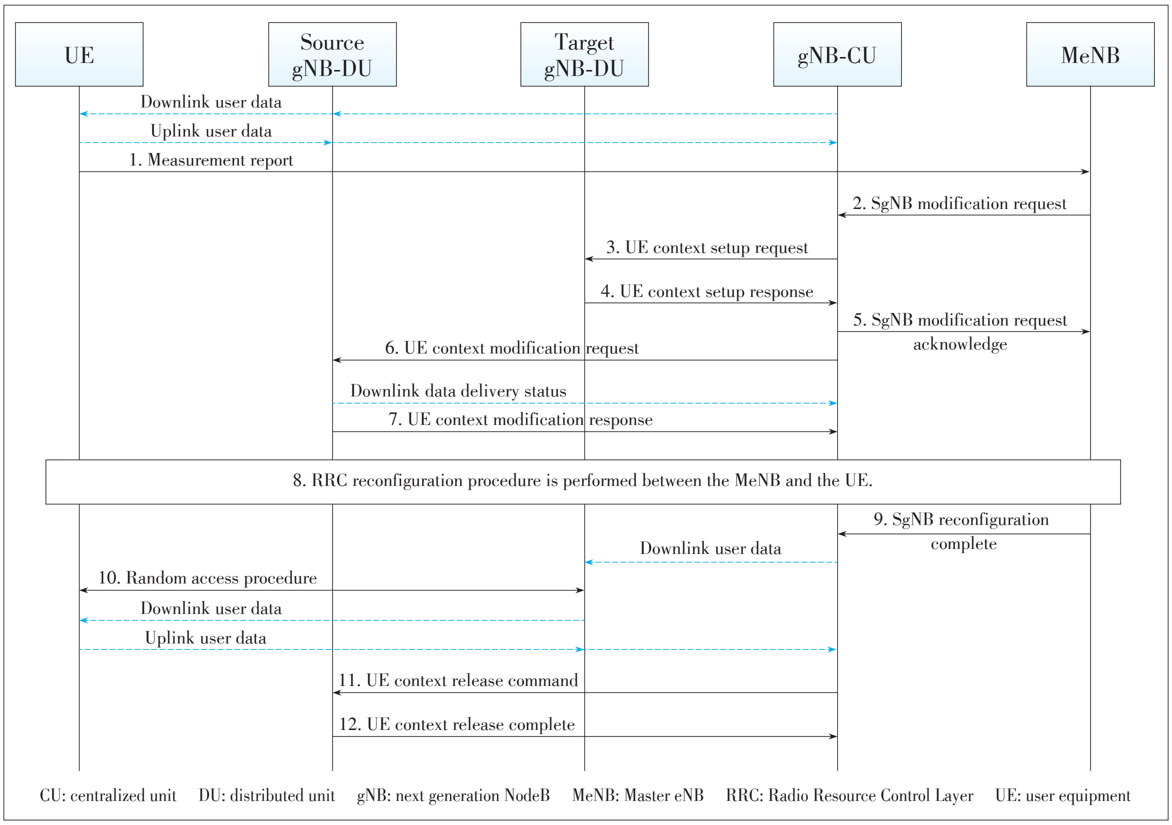
4) Option 7-3 (Only for DL)

Only the encoder resides in the LLS-CU, and the rest of PHY functions reside in the LLS-DU.

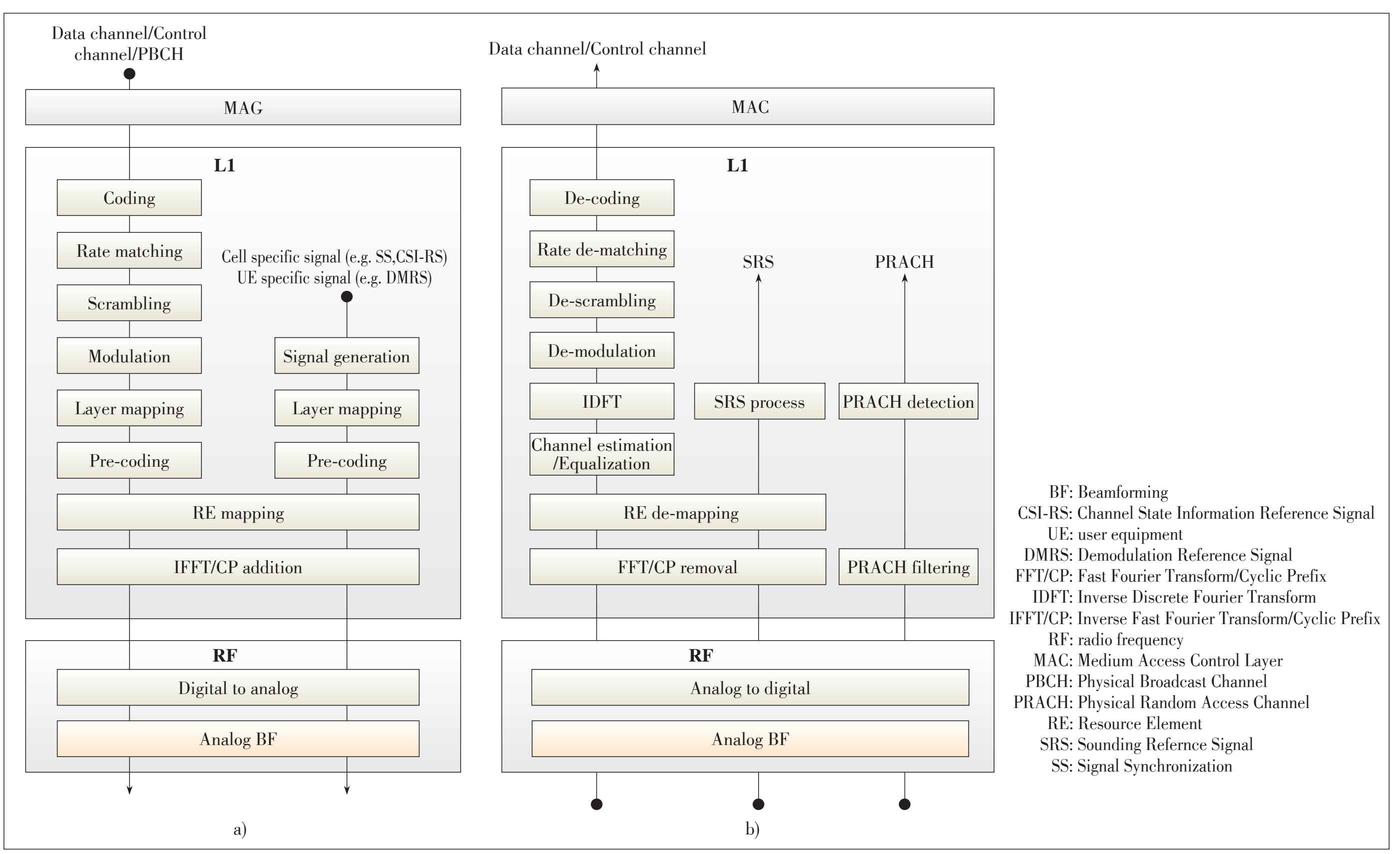
Additional potential functional split options were also considered. For the UL, there was a proposal to split between IDFT and Channel estimation/Equalization. Also, for both DL and UL, the possibility to split somewhere between Option 7-1 and Option 7-2 was proposed in light of digital beamforming [10].

6.2 Separation of Control Plane (CP) and User Plane (UP)

In order to provide the possibility of optimizing the location of different RAN functions based on the scenario and desired performance, the gNB-CU can be separated further into CU-CP



◀**Figure 5. Inter-gNB-DU mobility using MCG SRB in EN-DC [3].**



▲**Figure 6. One possible implementation of NR L1 processing chain at gNB for a) DL and b) UL [10].**

and CU-UP on the basis of HLS. The gNB-DU hosts the RLC/MAC/PHY protocols, the CU-CP hosts the control plane instance of PDCP and RRC protocols and the CU-UP hosts the user plane instance of PDCP (and SDAP) protocols [11]. The interface between CU-CP and CU-UP is named as E1. The overall RAN architecture with CU-CP and CU-UP separation is shown in **Fig. 7** [12].

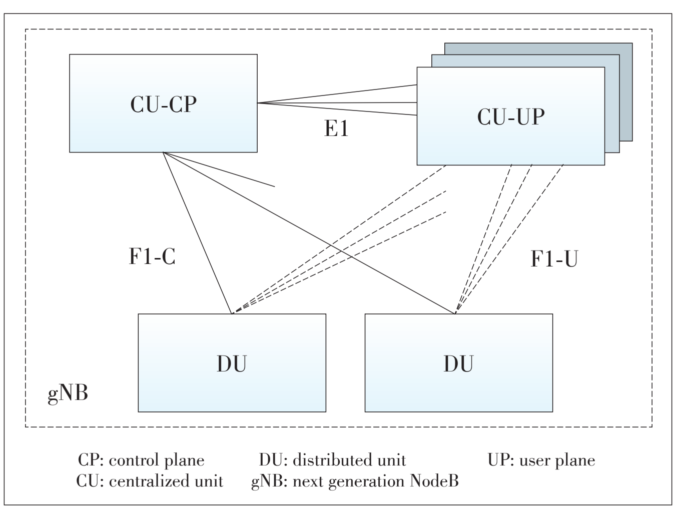
A gNB may consist of a CU-CP, multiple CU-UPs and multiple DUs. The CU-CP is connected to the DU through the F1-C interface, while the CU-UP is connected to the DU through the F1-U interface. The CU-UP is connected to the CU-CP through the E1 interface. Furthermore, one gNB-DU is connected to only one CU-CP and one CU-UP is connected to only one CU-CP. A gNB-DU or a CU-UP may be connected to multiple CU-CPs by appropriate implementation for resiliency. One gNB-DU can be connected to multiple CU-UPs under the control of the same CU-CP and one CU-UP can be connected to multiple gNB-DUs under the control of the same CU-CP.

The basic functions supported over the E1 interface include E1 interface management function and bearer management function, while such functions are still under investigation as E1 load management, E1 configuration update, inactivity detection, and new QFI notification.

6.3 CU-DU High Layer Split in E-UTRAN

In order to achieve better integration of LTE eNB with gNB, the converged architecture of LTE and NR is preferred, i.e. introducing central unit (LTE-CU) and distributed unit (LTE-DU) into Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) with PDCP/RLC split (option 2). This architecture aims to utilize the transport network in an efficient way and minimize the impacts on legacy LTE transport network. It is easier for further network upgrading when LTE CU and DU are deployed in operators' networks [13].

The CU-DU high layer split in E-UTRAN follows that in NR, including the function split architecture and interface function. Similar as gNB in NR, the eNB is split into two entities, i.e., eNB-CU and eNB-DU, and the interface between eNB-CU and eNB-DU is named as V1. The V1 interface supports the same functions as the F1 interface except that some LTE features depend on operators' requirements when the eNB is connected to EPC, such as NB-IoT and eMTC.



▲Figure 7. Overall RAN architecture with CU-CP and CU-UP separation [12].

7 Simulations

Based on the TCP throughput efficiency of data transmission, the simulations were conducted to show the performance of Options 2 and 3 (Fig. 1) for CU-DU split.

For Option 2, RRC and PDCP are in the central unit; RLC, MAC, physical layer and RF are in the distributed unit. For Option 3, low RLC (partial function of RLC, which mainly includes the segmentation related function), MAC, physical layer and RF are in distributed unit; RRC, PDCP and high RLC (the other partial function of RLC, which mainly includes the ARQ related function) are in the central unit.

For Option 3, since the Automatic Repeat Request (ARQ) is located in CU, the RLC retransmission suffers a two-way fronthaul delay, including the delay for RLC status report and the delay for the following data retransmission. Considering the mechanism of TCP, the delay of RLC retransmission may lead to some negative impact on the throughput.

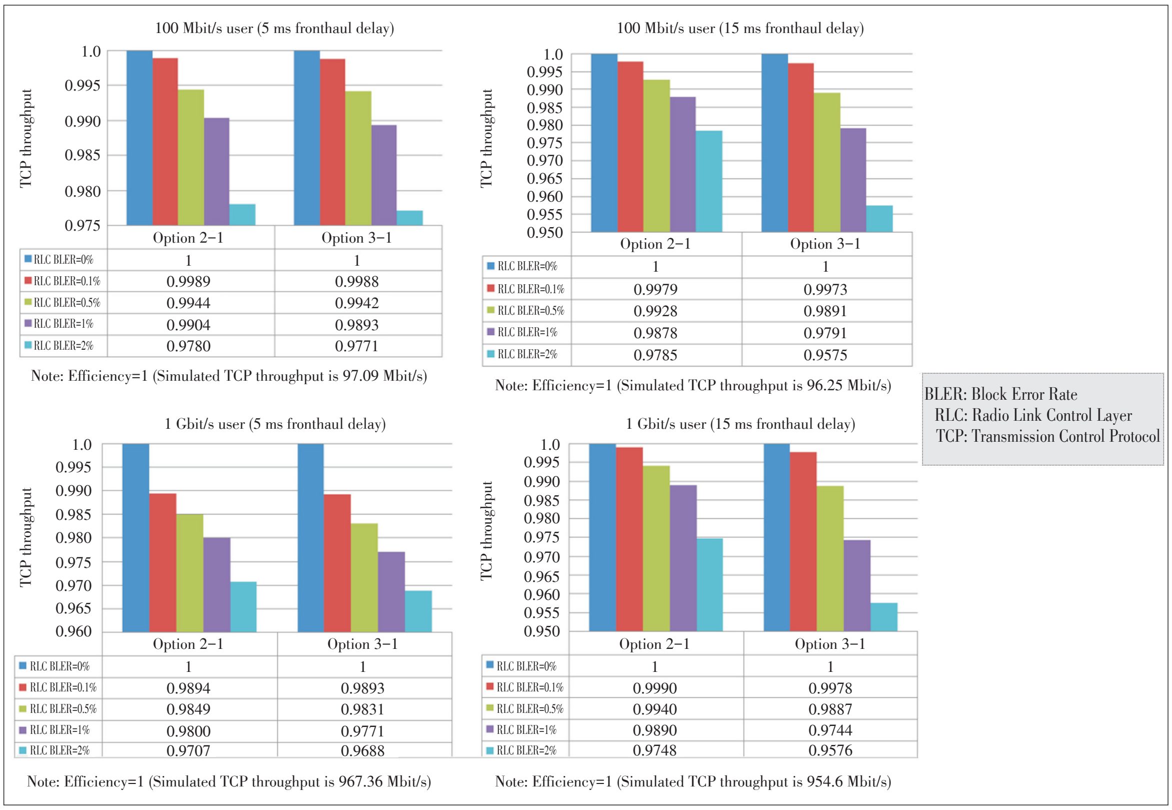
The simulation results of TCP throughput efficiency for Options 2 and 3 are shown in **Fig. 8** with different residual RLC BLER conditions. It can be observed that as the increase of the fronthaul delay, the TCP throughput will decrease. The TCP throughput of Option 3 decreases due to the additional retransmission delay from the fronthaul between CU and DU.

Slow-start is part of the congestion control strategy used by TCP. Once the slow-start threshold is reached, TCP changes from the slow-start algorithm to the linear growth (congestion avoidance) algorithm. Furthermore, if a ftp traffic model like 100 m file size and 1 G file size is used, the TCP slow start impacts performance more due to shorter simulation time. The simulation results in **Fig. 9** show that Option 2 performance is obviously better than Option 3 for short time TCP services considering TCP slow-start effects (initial TCP slow-start threshold set as 65,535).

It can be seen that Option 3 introduces extra RLC retransmission delay, and the extra delay may lead to negative impact on the throughput, especially for short time TCP services considering TCP slow-start effects. Compared with Option 3, Option2 provides better performance.

8 Conclusions

In this paper, we introduce the progress of CU-DU architecture and present the architecture for CU-DU split in NG-RAN. The CU-DU interface functions and basic mobility scenarios



100 Mbit/s user (5 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9989	0.9988
RLC BLER=0.5%	0.9944	0.9942
RLC BLER=1%	0.9904	0.9893
RLC BLER=2%	0.9780	0.9771

Note: Efficiency=1 (Simulated TCP throughput is 97.09 Mbit/s)

100 Mbit/s user (15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9979	0.9973
RLC BLER=0.5%	0.9928	0.9891
RLC BLER=1%	0.9878	0.9791
RLC BLER=2%	0.9785	0.9575

Note: Efficiency=1 (Simulated TCP throughput is 96.25 Mbit/s)

1 Gbit/s user (5 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9894	0.9893
RLC BLER=0.5%	0.9849	0.9831
RLC BLER=1%	0.9800	0.9771
RLC BLER=2%	0.9707	0.9688

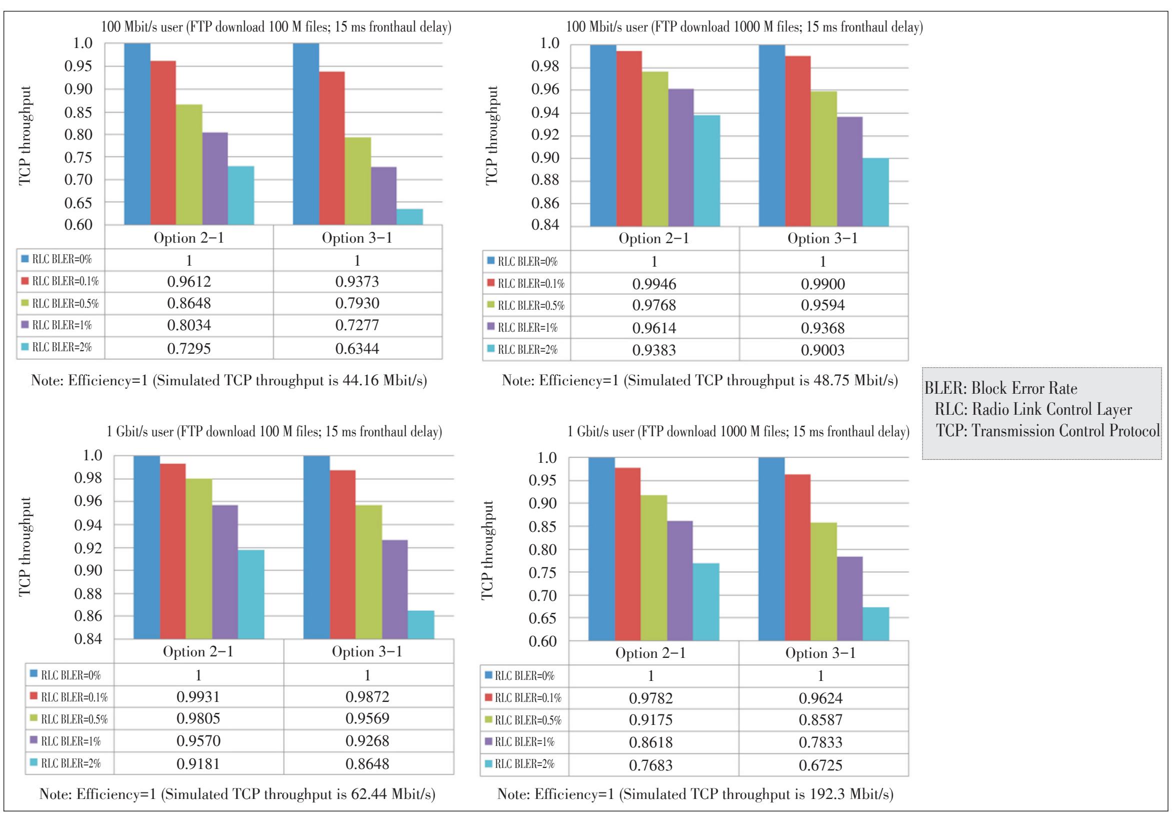
Note: Efficiency=1 (Simulated TCP throughput is 967.36 Mbit/s)

1 Gbit/s user (15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9990	0.9978
RLC BLER=0.5%	0.9940	0.9887
RLC BLER=1%	0.9890	0.9744
RLC BLER=2%	0.9748	0.9576

Note: Efficiency=1 (Simulated TCP throughput is 954.6 Mbit/s)

◀Figure 8. Simulation results on the data transmission in Options 2 and 3 (without TCP slow start impact).



100 Mbit/s user (FTP download 100 M files; 15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9612	0.9373
RLC BLER=0.5%	0.8648	0.7930
RLC BLER=1%	0.8034	0.7277
RLC BLER=2%	0.7295	0.6344

Note: Efficiency=1 (Simulated TCP throughput is 44.16 Mbit/s)

100 Mbit/s user (FTP download 1000 M files; 15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9946	0.9900
RLC BLER=0.5%	0.9768	0.9594
RLC BLER=1%	0.9614	0.9368
RLC BLER=2%	0.9383	0.9003

Note: Efficiency=1 (Simulated TCP throughput is 48.75 Mbit/s)

1 Gbit/s user (FTP download 100 M files; 15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9931	0.9872
RLC BLER=0.5%	0.9805	0.9569
RLC BLER=1%	0.9570	0.9268
RLC BLER=2%	0.9181	0.8648

Note: Efficiency=1 (Simulated TCP throughput is 62.44 Mbit/s)

1 Gbit/s user (FTP download 1000 M files; 15 ms fronthaul delay)

	Option 2-1	Option 3-1
RLC BLER=0%	1	1
RLC BLER=0.1%	0.9782	0.9624
RLC BLER=0.5%	0.9175	0.8587
RLC BLER=1%	0.8618	0.7833
RLC BLER=2%	0.7683	0.6725

Note: Efficiency=1 (Simulated TCP throughput is 192.3 Mbit/s)

◀Figure 9. Simulation results on the data transmission in options 2 and 3 (short time TCP services).

are discussed in this paper. The solutions to these challenges and potential optimization are also proposed. In addition, the other CU-DU related topics are also introduced, including CU-DU low layer split, separation of CP and UP, and the high layer split in E-UTRAN.

References

[1] NTT DOCOMO, INC., "Revised WID on new radio access technology," RP-172109, 2017.
[2] *Study on new Radio Access Technology: Radio Access Architecture and Interfaces*, 3GPP TR38.801, April. 2017.
[3] *NG-RAN, Architecture Description; (Release 15)*, 3GPP TS 38.401, Jan. 2018.
[4] *NG-RAN, F1 General Aspects and Principles; (Release 15)*, 3GPP TS 38.470, Jan. 2018.
[5] *NG-RAM, NR User Plane Protocol; (Release 15)*, 3GPP TS38.425, Dec. 2017.
[6] H. J. Zhang, N. Liu, X. L. Chu, et al., "Network slicing based 5g and future mobile networks: mobility, resource management, and challenges," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138–145, Aug. 2017. doi: 10.1109/MCOM.2017.1600940.
[7] H. J. Zhang, Y. Qiu, X. L. Chu, K. Long, and V. C. M. Leung, "Fog radio access networks: mobility management, interference mitigation and resource optimization," *IEEE Wireless Communications*, vol. 24, no. 6, pp. 120–127, Dec. 2017. doi: 10.1109/MWC.2017.1700007.
[8] ZTE, "Discussion on inter-DU mobility with dual connectivity," R3-180134, 2018.
[9] H. J. Zhang, C. X. Jiang, J. L. Cheng, and V. C. M. Leung, "Cooperative interference mitigation and handover management for heterogeneous cloud small cell networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 92–99, Jun. 2015. doi: 10.1109/MWC.2015.7143331.
[10] *Study of CU-DU Low Layer Split for NR (Release 15)*, 3GPP TS 38.816, Jan. 2018.
[11] Ericsson, New WID on separation of CP and UP for split option 2, RP-173831, 2017.
[12] *Study of Separation of NR Control Plane (CP) and User Plane (UP) for Split Option 2 (Release 15)*, 3GPP TS 38.806, Jan. 2018.
[13] China Unicom, Orange, China Telecom, Huawei, and HiSilicon, "Revised SID: study on eNB(s) architecture evolution for E-UTRAN and NG-RAN," RP-172707, Dec. 2017.

Manuscript received: 2018-02-05

Biographies

GAO Yin (gao.yin1@zte.com.cn) received the master's degree in circuit and system from Xidian University, China in 2005. Since 2005 she has been with the research center of ZTE Corporation and been engaged in the study of 3G/4G/5G technology. She has authored or co-authored about hundreds of proposals for 3GPP meetings and journal papers in wireless communications and has filed more than 100 patents. She was the rapporteurs of multiple 3GPP WIs. From August 2017, she has been elected as the vice chairman of 3GPP RAN3.

HAN Jiren (han.jiren@zte.com.cn) received the master's degree in wireless communication systems from University of Sheffield, UK in 2016. He is an advanced research engineer at the Algorithm Department, ZTE Corporation. His research interests include 5G wireless communications and signal processing.

LIU Zhuang (liu.zhuang2@zte.com.cn) received the master's degree in computer science from Xidian University, China in 2003. He is currently a senior 5G research engineer at ZTE R&D center, Shanghai. His research interests include 5G wireless communications and signal processing.

LIU Yang (liu.yang31@zte.com.cn) received the Ph.D. degree in communication and information systems from Beijing University of Posts and Telecommunications (BUPT), China in 2016. He was a visiting scholar at Department of Electrical and Computer Engineering of North Carolina State University, USA from 2013 to 2015. He is currently a 5G research engineer at ZTE R&D center, Shanghai. His research interests include 5G wireless communications and signal processing.

HUANG He (huang.he4@zte.com.cn) received the bachelor's degree in computer science and technology from Shanghai Jiao Tong University, China in 2004. He is currently the chief engineer of wireless innovation laboratory of ZTE Corporation and leads the research and standardization work on 5G RAN. He has filed more than 60 patents. He was the rapporteurs of multiple CCSA/3GPP SIs/WIs and the editors of the related protocols.