GTI

Sub-6GHz Indoor Solution and Small Cell

White Paper

http://www.gtigroup.org

GTI Sub-6GHz Small Cell

White Paper

Version:	V1.0
Deliverable Type	☐ **Procedural Document** √ **Working Document**
Confidential Level	☐ **Open to GTI Operator Members** ☐ **Open to GTI Partners** √ **Open to Public**
Working Group	Network
Task	Indoor Solution and Small Cell
Source members	Baicells, Sprint, CMCC, Datang, Nokia, Huawei, Ericsson, R&S
Support members	

Editor	Baicells, CMCC, Sprint, Datang
Last Edit Date	Jan 2019
Approval Date	

Confidentiality: This document may contain information that is confidential and access to this document is restricted to the persons listed in the Confidential Level. This document may not be used, disclosed or reproduced, in whole or in part, without the prior written authorization of GTI, and those so authorized may only use this document for the purpose consistent with the authorization. GTI disclaims any liability for the accuracy or completeness or timeliness of the information contained in this document. The information contained in this document may be subject to change without prior notice.

Document History

Date	Meeting #	Version #	Revision Contents
2019/1/9		V 1.0	The first version of GTI sub-6GHz indoor solution and small cell Whitepaper.

Table of Contents

1. Executive Summary

Mobile wireless communication has experienced explosive growth over the past decade. It can be foreseen that this trend of dramatic increase in traffic demand will continue over the next decade due to more and more additional wideband services and applications. The wireless industry has to take on the challenge to find some cost-effective solutions to 1000-fold increased traffic demand, which has become one of the main drives for network densification and capacity extension using small cells.

In the early years of small cell deployments, the chief priorities were to improve indoor coverage, initially in the home; to fill ad hoc gaps in outdoor coverage; and to supplement capacity in a localized way via urban or indoor hotspots.

Nowadays, 5G mobile networks will use Sub-6 GHz, such as n 41, n77, n78, n79 and mmWave that are on much higher frequency bands than 2G, 3G, and 4G. High-band 5G signals will have greater link loss when penetrating through walls. The mmWave band has even greater link loss, making signals on this band even more difficult to travel through any walls. Therefore having access points at closer proximity to end user becomes much more crucial for mobile network. This implies that small cells are ideally suited for delivering many 5G network solutions, providing not only coverage and capacity, but also lower latency and higher quality of service in lower cost.

Along with the emerging and mature of SDN/NFV technology, SDN/NFV based 5G mobile network provides more flexibility and scalability for small cell deployment. Also CU/DU split and distributed RRU architecture make small cells more versatile.

2. Abbreviations

Abbreviation	**Explanation**
2/3/4/5G	The 2nd/3rd/4th/5th Generation Telecommunication
3GPP	The 3rd Generation Partnership Project
AR	Augmented Reality
BLE	Bluetooth Low Energy
CAPEX	Capital Expense
CPRI	Common Public Radio Interface
C-RAN	Centralized/Cloud RAN
CSI	Channel State Information
CU	Central Unit
DAS	Distributed Antenna System
DMRS	Demodulation Reference Signal
DU	Distributed Unit
eMBB	Enhanced Mobile Broadband
gNB	NR node
GW	Gateway
MAC	Medium Access Control
MIMO	Multiple-Input Multiple-Output
HD	High Definition
mMTC	Massive Machine Type Communication
mmWave	Millimeter Wave
MU-MIMO	Multi-User MIMO

NFV	Network Function Virtualization
NR	New Radio
O&M	Operation and Maintenance
OAM	Operation, Administration, and Maintenance
OFDM	Orthogonal Frequency Division Multiplexing
OPEX	Operational Expense
PHY	Physical Layer
RE	Radio Equipment
REC	Radio Equipment Control
RLC	Radio Link Control
RRC	Radio Resource Control
RRU	Remote Radio Unit
SDN	Software Defined Network
SRS	Sounding Reference Signal
SU/MU	Single User/Multi-User
UE	User Equipment
UL	Uplink
UP	User Plane
URLLC	Ultra-Reliable and Low Latency Communications
VR	Virtual Reality

3. Introduction

5G era

The 5G era will mean a connected society, in which people, machines and things are always connected in the most appropriate manner, providing a variety of services and offering a number of diverse business models. Examples include smart cities, intelligent transportation systems, autonomous cars, the Internet of things, tactile Internet and new worlds of immersive user experience based on augmented/virtual/mixed realities.

5G use cases

There are three main use cases enabled by eMBB (enhanced mobile broadband), mIoT (massive Internet of things) and URLLC (ultra-reliable and low-latency communications).

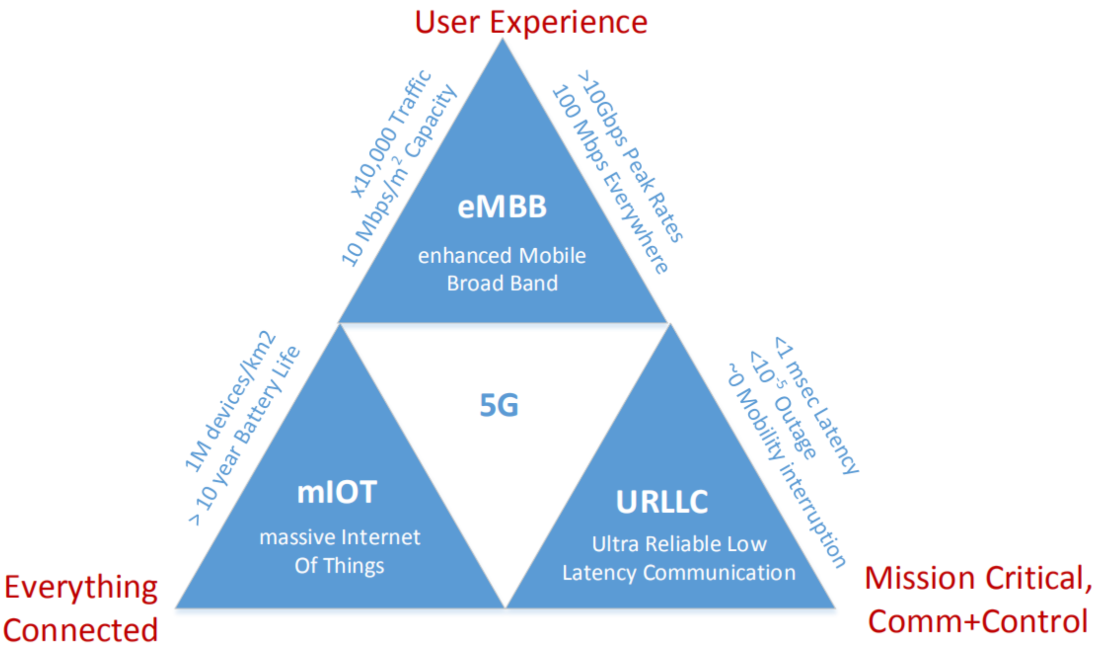


Figure 3-1 5G use cases

Small Cell

What is the role of small cells in the realization of 5G? Many 5G use cases, such as eMBB and LLC, are undoubtedly best realized by radio stations and compute/storage resources that are close to end users or devices. For example, close proximity enables higher signal quality and hence higher data rates, lower latency on the radio network, as well as lower end-to-end latency if edge computing is also deployed. This implies that small cells are ideally suited for delivering many 5G network solutions.

Indoor Scenario

Statistics show that more than 80% services on 4G mobile networks occur indoors. The industry predicts that a greater number of mobile services will take place indoors as 5G spurs diverse services and extends business boundaries. Therefore, indoor mobile networks of the 5G era will become an integral part of operators' core competitiveness.

4. References

[1] 3GPP, TR 36.932, Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN, v15.0.0

[2] 3GPP, TS 23.501, NR; System Architecture for the 5G System Stage-2, v15.0.0

[3] 3GPP, TR 36.842, Study on Small Cell enhancements for E-UTRA and E-UTRAN; Higher layer aspects

[4] 3GPP, TR 36. 872, Study on Small Cell enhancements for E-UTRA and E-UTRAN; Physical layer aspects

[5] NGMN 5G white paper

[6] Indoor 5G Network Whitepaper, Huawei GSA HKT

[7] eCPRI Interface Specification

[8] IEEE 1914 NGFI (xhaul): efficient and scalable fronthaul transport for 5G

[9] ITU-T technical report: Transport network support of IMT-2020/5G

[10] 3GPP, TR38.874 v1.0.0, Study on Integrated Access and Backhaul

[11] 3GPP, TS38.214 v15.3.0, NR, Physical Layer Procedures for Data, Release 15

[12] Control, User and Synchronization Plane Specification xRAN Fronthaul Working Group XRAN-FH.CUS.0-v02.00

5. Scenarios

In addition to supporting the evolution of the established prominent mobile broadband use cases, 5G will support countless emerging use cases with a high variety of applications and variability of their performance attributes: From delay-sensitive video applications to ultra-low latency, from high speed entertainment applications to mobility on demand for connected objects, from best effort delivered across a wide range of devices (e.g. smart phones, wearable, MTC) to across a fully heterogeneous environment.

Below we will discuss some indoor scenarios covering eMBB, URLLC and mMTC features.

5.1. Smart Office

In a future office, it is envisioned that most of the devices will be wireless connected. Users will interact through multiple and wireless connected devices, such as VR, AR, Telepresence etc.. This suggests a scenario in which hundreds of users require ultra-high bandwidth for services that need high-speed execution of bandwidth-intensive applications. Ultra-high traffic volume, and for some application latency, are the main challenges applicable for this use cases.

However, the need of ultra-high bandwidth may not be constant from daytime to night. Due to mobile users (office workers) mobility, the traffic fluctuates a lot between in the daytime and during the night. People gathering and ultra-high bandwidth required only occur during office hours, while in the mid-night, much more spare resources can be shared. So it is necessary to deploy a network which can share resources among different scenarios accordingly and automatically for efficiency.

5.2. Large Venue

This use case is characterized by a high connection density and potentially temporary use (e.g., in a stadium, concert, or other conference events). Several hundreds of thousands of users may be served per km^2, possibly integrating physical and virtual information such as score, information on athletes or musicians, etc., during the event. People can watch the game using augmented reality (AR) or virtual reality (VR), share high definition (HD) live video or post HD photos to social networks. These applications will require a combination of ultra-high connection density, high data rate and low latency. Meanwhile, in order to support the burst of the fluctuating traffic data, a network with elastic capacity should be prepared.

5.3. Hospital/eHealth

In many scenarios, specialists or medical expertise are not always available in many hospitals and could not join a local surgeon remotely to perform certain procedures which require expert skills. It is expected that in the near future the physicians could be able to command a tele-robot at the patient’s location, allowing remote physical examination with fully AV and haptic feedback via wireless network. Also, in a robotics assisted tele-surgery scenario, the location of the specialists could never be the confined factor for operating expert procedures remotely through wireless network. To make these ehealth scenarios realistic, the network performance must achieve both ultra-reliability and ultra-low latency as well as high bandwidth.

5.4. Smart Factory

In smart factory, automation is a key feature which heavily relies on the use of robots and machine intelligence. Automation will complement human workers, not only in jobs with repetitive tasks (e.g., production, transportation, logistics, office/administrative support) but also with high-precision operations. In order to enable these application with completely diverse tasks, it will be essential to provide an underlying control network with very low latency and high reliability. For many robotics scenarios in manufacturing a round-trip reaction time of less than 1ms is anticipated.

5.5. Smart Warehouse

Smart warehouse heavily depends on automation and real-time connection. The most obvious aspect of a smart warehouse is the real-time connection of thousands of things in the warehouse. When you have a smart warehouse, you should be able to get real-time updates on all the activity happening day or night.

In addition, smart warehouse, e.g. logistics hub, is also expected to have the ability of connecting massive amounts of goods via wireless network for efficient management, dispatch, and transshipment. A wide range of goods need to be identified, positioned, sorted, and then automatically transported to dedicated locations with the help of robots. It is estimated that one object needs to be connected for every square meter.

6. Challenges and Requirements

6.1. Deployment Related

6.1.1. Flexibility

The more people have been able to achieve while on the move, the more dependent society has become on mobile broadband networks, and then the higher speed the new services and features emerge in. Keeping pace with ever-increasing demand calls for much more network functions, 5G network need to deliver the need of much higher level of flexibility in order to avoid second deployment.

6.1.2. Lower Cost Resolution

The increasing use of smart mobile devices has escalated the competition among mobile operators, and the competition has driven them to upgrade bandwidth and improve quality of services, while average revenue and profitability per user are declining. This unfavorable cycle has witnessed the obstacle for the capacity of conventional RAN infrastructures as it has been challenged by the increases of volumes of data and the urged demands for high-speed communications. In fact, conventional RAN platforms are proprietary deployments, and the upgrades of hardware interfaces would mean significant expenditure for mobile service operators even when profitability is declined.

In addition the need for small cells will be even more critical in 5G networks due to the introduction of higher spectrum bands, which necessitate denser network deployments to support larger traffic volumes per unit are.

In the light of foregoing, denser deployments make higher cost, therefore, cost-effective small cell solution is crucial for operators.

6.1.3. Fronthaul of C-RAN

4G LTE networks have started moving towards a distributed Cloud or Centralized RAN (C-RAN) based architecture. In 5G, C-RAN architecture will play a dominate role in facilitating CoMP (Coordinated Multipoint Processing) for purpose of mitigating inter-cell interference and increasing spectral efficiency. In addition, C-RAN is one of the key enablers in reducing costs and providing deployment flexibility during significant network densification.

The C-RAN approach advocates for the separation of the radio elements of the base station from the elements processing the baseband signal, which are centralized in a single location or even virtualized into the cloud. This approach benefits from simpler radio equipment at the network edge, easier operation, and cheaper maintenance, while the main RAN intelligence is centralized in the operator-controlled premises. However in this architecture, the challenge is that such a functional split requires these elements to be connected through a high-speed, low-latency, and accurately synchronized network, the so-called fronthaul.

6.1.4. Backhaul/Midhaul

During small cell deployment, backhaul is usually the bottleneck in achieving the potential capacity gain. Network densification won’t be able to have the expected performance improvement unless it is complemented by densification of the backhaul that connects macro and small cells to the core network. Ideally availability of fibre-based backhaul in dense urban environments will offer the opportunity of using C-RAN architecture and CoMP in small deployment. Unfortunately, it is very unlikely to fibre based backhaul available for all the small cells due to high costs and other location limitations. Providing high capacity backhaul for small cells becomes one of the biggest challenges in 5G network densification. At locations without fibre based backhaul, wireless backhaul and Integrated Access and Backhaul (IAB) will have to be used, which will be discussed in further details in next section.

6.2. Services Related

6.2.1. Coverage

5G mobile networks will use Sub-6GHz and mmWave that are on much higher frequency bands than 2G, 3G, and 4G. High-band 5G signals will have greater link loss when penetrating through walls. The mmWave band has even greater link loss, making signals on this band even more difficult to travel through any walls. As a result, it is growing extremely difficult to provide sound outdoor-to-indoor coverage.

Low-powered indoor small cells are needed to offer a simple and cost effective solution to provide indoor coverage to complement macro cell coverage in order to enhance user experience.

6.2.2. Elastic Capacity

Data traffic tends to peak when crowds aggregate in large venues, entertainment centers, and other densely populated areas. When it comes to scenarios where companies cluster together (industrial parks, office buildings, and entrepreneurship innovation parks), explosive capacity requirements emerge during office hours or commercial events. Therefore, in order to keep up with the booming and diversified 5G services, a network with elastic capacity ought to be adequately prepared. Such network not only need to meet the changing requirements of services volume as time and areas vary, but also must cope with rapid surges in data traffic.

Also, network capacity must feature redundancy and scalability. Only then can the network respond to the real-time capacity demand in hotspot areas and carry versatile services.

6.2.3. Low Latency

In 5G, services featuring the Tactile Internet, high-resolution video streaming, tele-medicine, tele-surgery and real-time control dictate new specifications for low latency. Latency critical services in 5G networks demand an E2E delay of 1ms to 100ms. Some use cases such as VR and online gaming may require round trip latency on the order of 1 ms.

6.2.4. Reliability

Industry applications such as smart manufacturing and telemedicine depend on precise control, which makes reliability a demanding issue. According to 3GPP, the network reliability must be higher than 99.999%. To get a feel of the low tolerance to error in 5G, here are some comparative values: LTE tolerates Block Error Ratio of 0.01, and the BLER for 5G is expected to be 0.00001 in a 1 ms period.

Meanwhile, the CSP (Communication Services Provider) would also require an assurance system to assure such high levels of reliability at all points of time, e.g., under massive IoT load conditions, and in dynamic resource allocation situations.

6.2.5. Massive Connectivity

In the 5G era, Internet of Things (IoT) will become reality which leads to enormous number of ultra-high-density wireless connections, such as smart warehouse. Therefore it is basic and critical for the network to have the capability of connecting massive devices and providing services at a time.

7. Architecture and Technology

7.1. 5G Architecture

The 3GPP's architecture of the 5G network is shown below in Figure 7-1. The NG-RAN consists of a set of gNBs connected to the 5GC (5G Core Network) via the NG interface. The gNBs can be interconnected through the Xn interface. A gNB may consist of gNB-Centralized Unit (gNB-CU) and gNB-Distributed Unit (gNB-DU). The CU processes non-real time protocols and services, and the DU processes PHY level protocol and real time services. A gNB-CU and the gNB-DU units are connected via F1 logical interface.

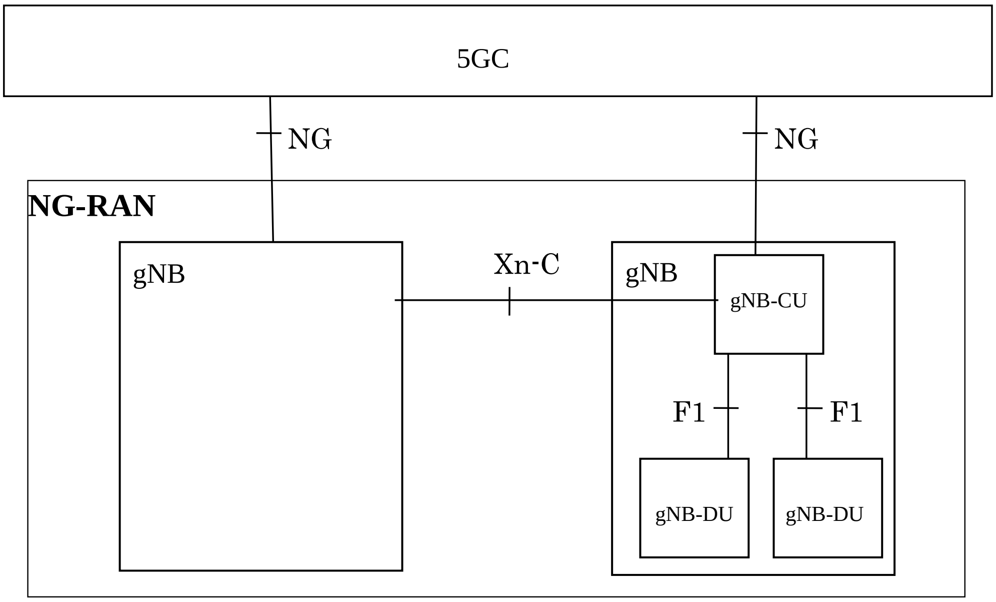


Figure 7-1 5G Architecture

The gNB can also be split into three parts: CU, DU and RRU (Remote Radio Unit), Where fronthaul is the network between RRU and DU (CPRI/eCPRI/NGFI interfaces), midhaul is the network between DU and CU (F1 interface) and backhaul is the network between CU and 5G CN (NG interface) and between CUs (Xn interface). The fronthaul would typically be based on LL FS (low layer function split) and the midhaul would typically be based on HL FS (high layer function split).

Evolving from 4G/LTE to 5G New Radio (NR) transport architecture, the main change is that the original BBU function in 4G/LTE is split into three parts: CU, DU, and RRU. The motivation of this design is manifold, for example, the new design could better facilitate radio access network (RAN) virtualization.

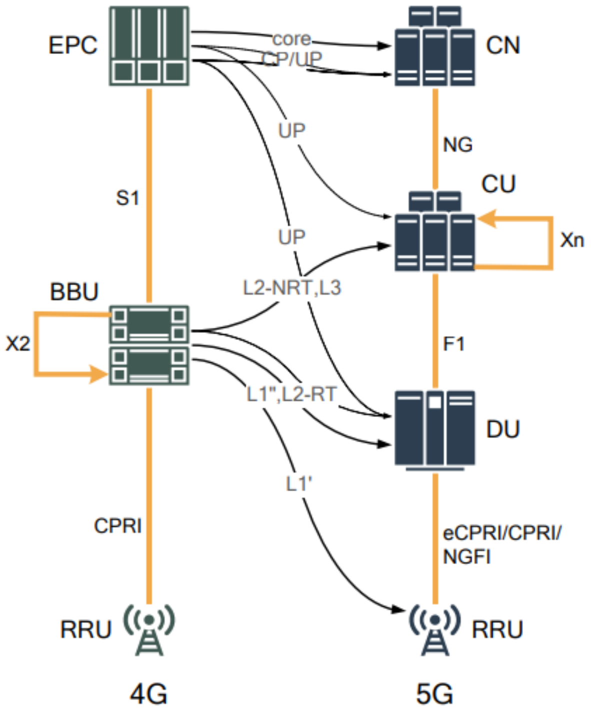


Figure 7-2 4G and 5G architecture relationship

As a part of study item for NR, 3GPP started studying different functional splits between central and distributed units. They have proposed 8 possible options shown in below figure.

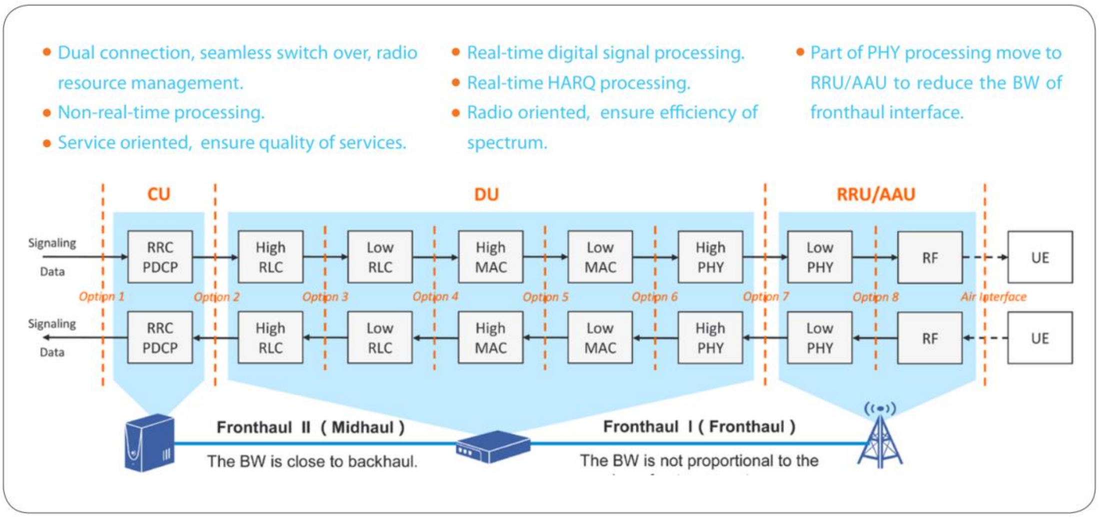


Figure 7-3 Function split options

- Option 1: RRC/PDCP split;

- Option 2: PDCP/RLC split;
- Option 3: Intra-RLC split-High RLC and Low RLC;
- Option 4: RLC/MAC split;
- Option 5: Intra-MAC split-High MAC and Low MAC;
- Option 6: MAC/PHY split;
- Option 7: Intra-PHY split-High PHY and Low PHY;
- Option 8: PHY/RF split;

Some of the benefits of an architecture with the deployment flexibility to split and move NR functions between central and distributed units are below:

- Flexible HW implementation allows for scalable cost-effective solutions;
- A split architecture allows for coordination for performance features, load management, real-time performance optimization, and enables NFV/SDN.
- Configurable functional split enables adaption to various use cases, such as variable latency and bandwidth.

In the following context, we will only introduce Option 2 and Option 7 split options.

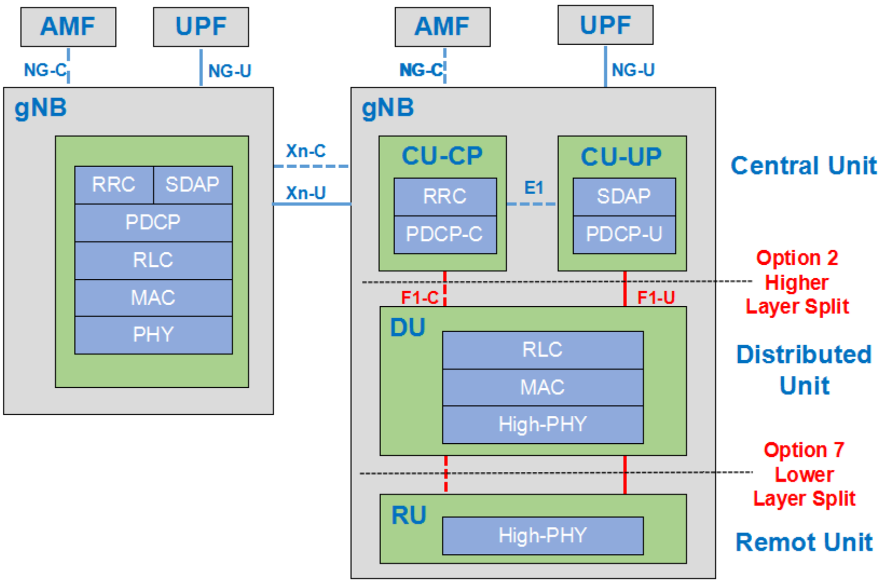


Figure 7-4 Split Option 2 and Option 7 Interface

7.1.1. CU/DU Split

Option 2: In this split option, RRC,PDCP are in the central unit. RLC, MAC, PHY and RF are in the distributed unit. In addition, this option can be achieved by separating the RRC and PDCP for the CP stack and the PDCP for the UP stack into different central entities.

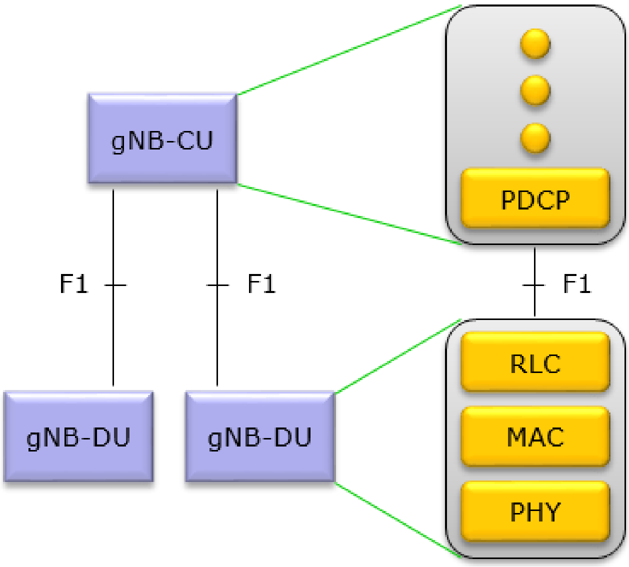


Figure 7-5 CU/DU Option 2 Split

7.1.2. DU/RU Split

The functional split point of RU and DU is Option 7, there are multiple possible realizations of Option 7, including asymmetrical options which allow obtaining benefits of different sub-options for UL and DL independently.

Option 7-1 in this option the UL, FFT, CP removal and possibly PRACH filtering functions reside in the RU, the rest of PHY functions reside in the DU. In the DL, iFFT and CP addition functions reside in the RU, the rest of PHY functions reside in the DU.

Option 7-2 in this option the UL, FFT, CP removal, resource de-mapping and possibly pre-filtering functions reside in the RU, the rest of PHY functions reside in the DU. In the DL, iFFT, CP addition, resource mapping and precoding functions reside in the RU, the rest of PHY functions reside in the DU.

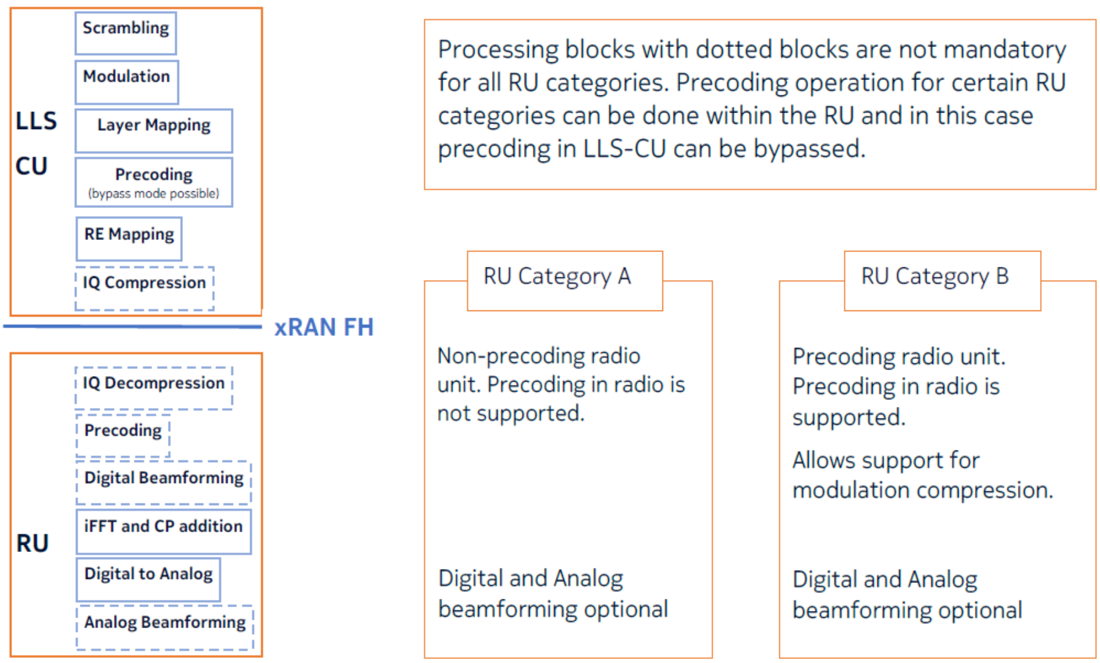
Option 7-3: Only the encoder resides in the DU, and the rest of PHY functions reside in the RU.

When considering the functional split defining a fronthaul interface there are two competing interests:

a) There is a benefit in keeping an RU as simple as possible because size, weight, and power draw are primary deciding considerations and the more complex an RU, the larger, heavier and more power-hungry the RU tends to be;

b) There is a benefit in having the interface at a higher level which tends to reduce the interface throughput relative to a lower-level interface – but the higher-level the interface, the more complex the RU tends to be.

To resolve this conundrum, xRAN has selected a single split point, known as “7-2x” but allows a variation, with the precoding function to be located either “above” the interface in the lls-CU or “below” the interface in the RU. For the most part the interface is not affected by this decision, but there are some impacts namely to provide the necessary information to the RU to execute the precoding operation. RUs within which the precoding is not done (therefore of lower complexity) are called “Category A” RUs while RUs within which the precoding is done are called “Category B” RUs. See Figure 7-6 for a depiction of this dual-RU concept.



Processing blocks with dotted blocks are not mandatory for all RU categories. Precoding operation for certain RU categories can be done within the RU and in this case precoding in LLS-CU can be bypassed.

RU Category A: Non-precoding radio unit. Precoding in radio is not supported.

Digital and Analog beamforming optional

RU Category B: Precoding radio unit. Precoding in radio is supported.

Allows support for modulation compression.

Digital and Analog beamforming optional

Figure 7-6 Split Point and Category A and Category B Radio Units

DL functional split for various physical layer channels is illustrated in Figure 7-7 (NR Category A RUs), and Figure 7-8 (NR Category B RUs).

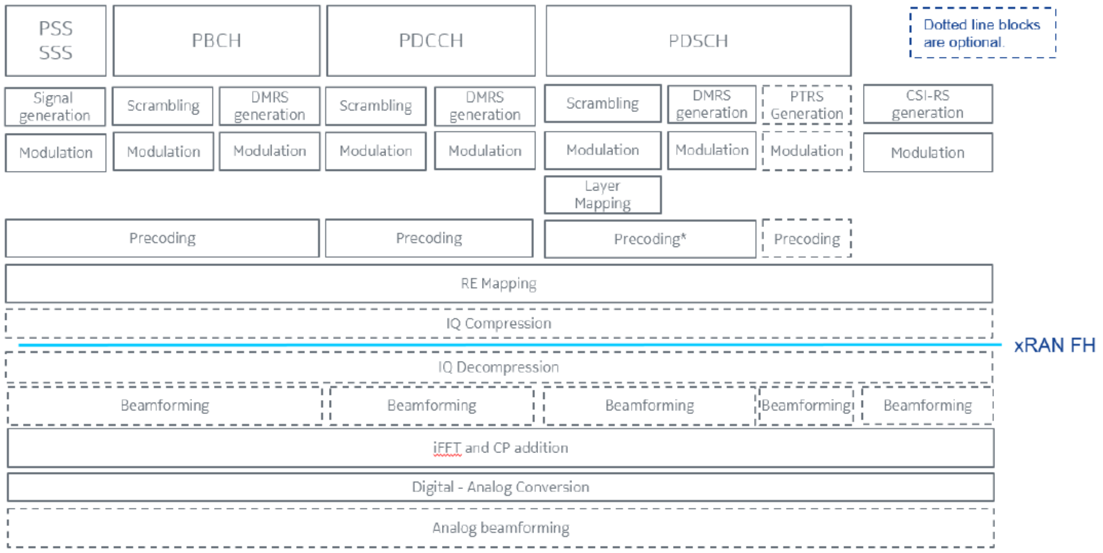


Figure 7-7: Lower layer DL split description, NR, Category A Radio

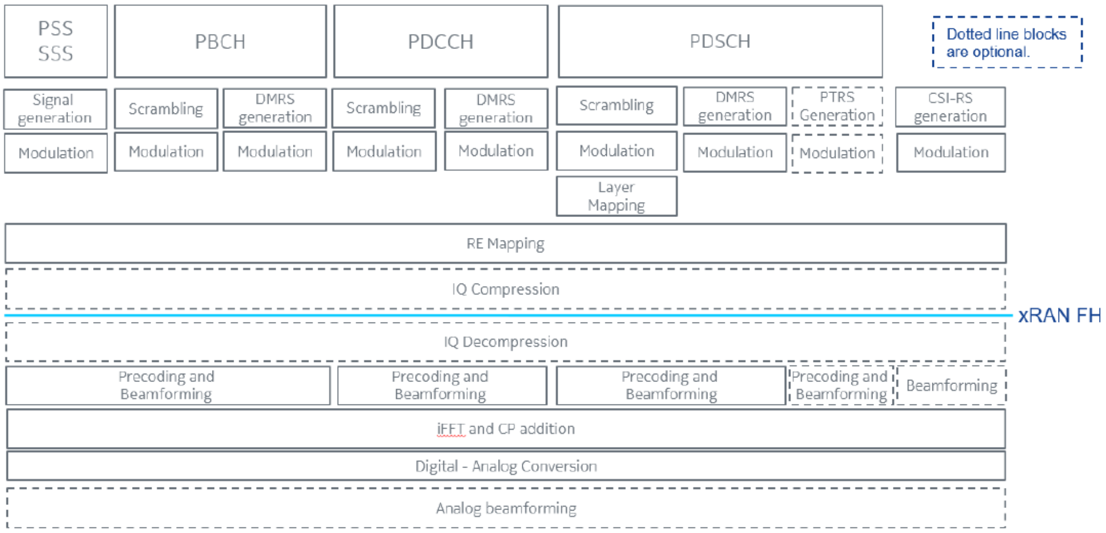


Figure 7-8: Lower layer DL split description, NR, Category B Radio

UL functional split for various physical layer channels and transmission modes are illustrated in Figure 7-9.

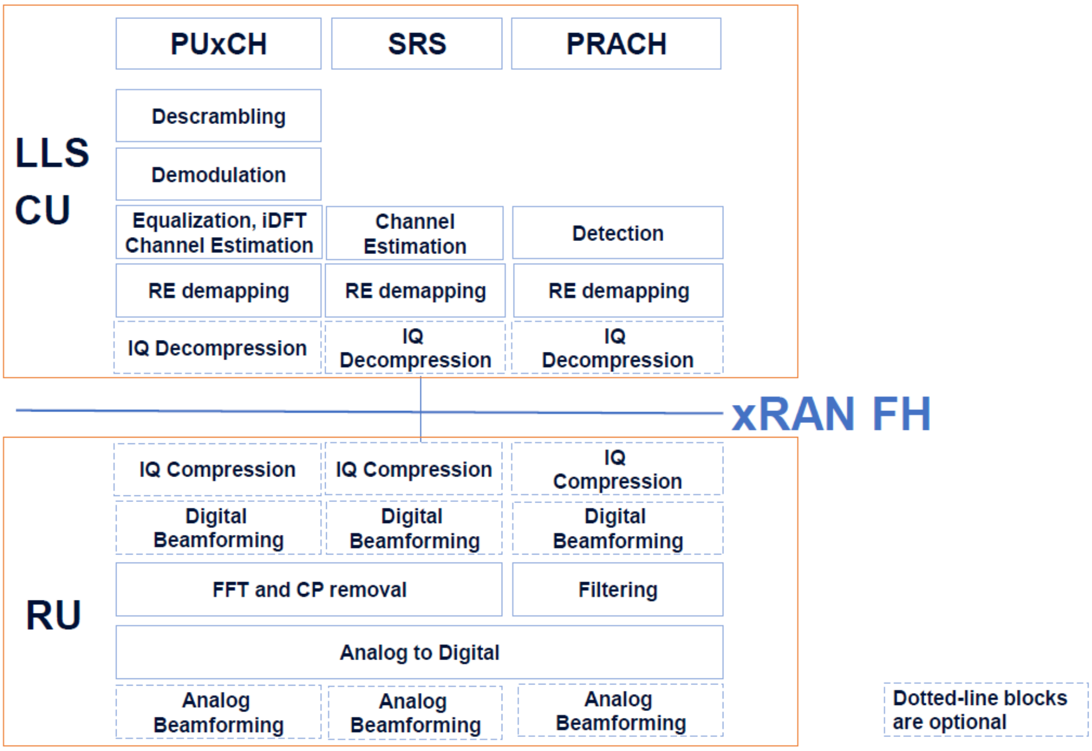


Figure 7-9 : Lower layer UL split description

7.2. SDN & NFV

7.2.1. Software Defined Network

Software-Defined Network (SDN) technology is an approach to cloud computing that facilitates network management and enables programmatically efficient network configuration in order to improve network performance and monitoring. SDN is meant to address the fact that the static architecture of traditional network is decentralized and complex while current network require more flexibility and easy troubleshooting. The control plane consists of one or more controllers which are considered as the brain of SDN network where the whole intelligence is incorporated.

SDN architecture decouple network control and forwarding functions, enabling network control to become directly programmable and the underlying infrastructure to be abstracted from applications and network services. This architecture includes:

- Directly programmable: Network control is directly programmable because it is decoupled from forwarding function.

- Agile: Abstracting control from forwarding lets administrators dynamically adjust network-wide traffic flow to meet changing needs.

- Centrally managed: Network intelligence is (logically) centralized in software-based SDN controllers that maintain a global view of the network, which appears to applications and policy engines as a single, logical switch.

- Programmatically configured: SDN lets network managers configure, manage, secure, and optimize network resources very quickly via dynamic, automated SDN programs, which they can write themselves because the programs do not depend on proprietary software.

- Open standard-based and vendor-neutral: When implemented through open standards, SDN simplifies network design and operation because instructions are provided by SDN controllers instead of multiple, vendor-specific devices and protocols.

7.2.2. Network Function Virtualization

Network function virtualization (NFV) is a network architecture concept that uses the technologies of IT virtualization to virtualize entire classes of network node functions into building blocks that may connect, or chain together, to create communication services.

The NFV framework consists of three main components:

- Virtualized network functions (VNFs) are software implementations of network functions that can be deployed on a network functions virtualization infrastructure (NFVI)
- Network functions virtualization infrastructure (NFVI) is the totally of all hardware and software components that build the environment where the VNFs are deployed. The NFV infrastructure can span several locations. The network providing connectivity between these locations is considered as part of the NFV infrastructure.
- Network functions virtualization management and orchestration architectural framework (NFV-MANO) is the collection of all functional blocks, data repositories used by these blocks, and reference points and interfaces through which these functional blocks exchange information for the purpose of managing and orchestrating NFVI and VNFs.

7.2.3. Relationship between SDN and NFV

SDN is a concept related to NFV, but they refer to different domains. NFV could efficiently complement the SDN functions.

In essence, SDN is an approach to build data networking equipment and software that separates and abstracts elements of these systems. It does this by decoupling control plane and data plane from each other, such that the control plane resides centrally and the forwarding components remain distributed.

NFV is not dependent on SDN or SDN concepts. It is entirely possible to implement a virtualized network function as a standalone entity using existing network and orchestration paradigms. However, there are inherent benefits in leveraging SDN concepts to implement and manage an NFV infrastructure, particularly when looking at the management and orchestration of VNFs.

An NFV infrastructure needs a central orchestration and management system that takes operator requests with a VNF, translates them into the appropriate processing, storage and network configuration needed to bring the VNF into operation. Once in operation, the VNF potentially must be monitored for capability and utilization, and adapted if necessary.

All these functions can be accomplished using SDN concepts and NFV could be considered one of the primary SDN use cases in service provider environments. It is also apparent that many SDN use-cases could incorporate concepts introduced in the NFV initiative. Example include where the centralized controller is controlling a distributed forwarding function that could in fact be also virtualized on existing processing or routing equipment.

7.3. Cell Split and Cell Merge

It is common that the capacity demand of a network varies from time to time during the course of a single day. In this case, dense deployments mean a waste of resources and result in stronger interference. With cell split and cell merge, the number of cells can be adaptively arranged based on the capacity demand of the users. Meanwhile, accurate coverage can be ensured across the entire cell. For example, in a large restaurant, operators can reduce cells in working hours and increase them in dining time to thus fully deliver the capacity needs with a minimal amount of capital expenditure (CAPEX)

7.4. Interference Cancellation

Some options of interference cancellation mechanisms include:

- Using C-RAN architecture and CoMP: Under C-RAN architecture, either PHY or MAC layer processing of base stations can be centralized so that Interference can be mitigated through coordinated processing (CoMP), such as coordinated scheduling, coordinated beamforming and joint transmission / receiving. CD/DU lower split and front haul with low latency and high bandwidth will have to be supported.

- Interference shaping: Arrange non-overlapping resource allocation areas in frequency domain for neighbouring cells to mitigate co-channel interference. The non-overlapping resource allocation areas are defined based on traffic pattern of the neighbouring cells. X2 interfaces between macro cell and small cells are required.

- Interference aware UL scheduling: UL interference can be reduced by scheduling separate PRB resource areas for UEs in adjacent cells. Base station uses filtered noise and interference measurements as well as UE transmit power density to identify whether the UE is at cell centre, cell middle or cell edge. Cell centre UEs that cause low interference will be allocated at the preferred scheduling area to achieve the best performance, while cell edge UEs will be allocated in other areas to reduce the interference.

- Interference aware UL power control: UL interference can be also reduced by UL power control. This will only work when this mechanism is implemented in all the cells in the area. Base station estimates UL interference based on UL SINR measurement. If the UL SINR is too low, the base station will send a command to the UEs in the cell to reduce the UL transmit power to reduce the interference to other cells. No X2 interfaces between macro cells and small cells are required.

- eICIC (enhanced inter-cell interference cancellation): DL interference from macro cell can be also mitigated through almost blanking subframe (ABS). Macro cell does not transmit any data in predefined ABS subframes, which allows small cells allocate cell edge UEs in those ABS subframes. This way, the strongest interference from macro cells to cell edge users can be greatly reduced.

- Interference nulling through beamforming: Transmit beamforming can be implemented in both macro cells and small cells to maximize the SINR of the served UE. The interference from neighbouring cell is nulled through linearly combining the beamforming weights based on UE feedback.

7.5. Common Public Radio Interface

CPRI defines key interface specification between REC (Radio Equipment Control) and RE (Radio Equipment) of radio base stations. CPRI is the short form of Common Public Radio Interface. CPRI is popular standard for transporting baseband I/Q signals to the radio unit in traditional BS (Base Station). CPRI allows efficient and flexible I/Q data interface for various standards e.g. GSM, WCDMA, LTE etc.

Although CPRI has been the main Fronthaul interface standard, many operators started to question its suitability to high bandwidth 5G use cases. Improvements to efficiency and link capacity utilization were requested and also advanced networking and OAM features of mainstream packet transport standards were requested. In the light of the previous statement, eCPRI was defined to meet these requirements. eCPRI standard defines specification which connects eREC and eRE via fronthaul transport network e.g. 5G, LTE-advanced Pro etc.

eCPRI is packet based fronthaul interface developed by CPRI forum. It has the same level of interoperability as CPRI. It designed to handle diverse fronthaul types. eCPRI interface includes the following information flows:

- User plane

  - User Data: User information (to be transmitted from/to the base station to/from the user equipment) with format depending on the underlying functional decomposition between the eREC and the eRE.

  - Real-Time Control data: Time-critical control and management information directly related to the User Data.

  - Other eCPRI services: eCPRI services such as User Plane support, remote reset, etc.

- Control and Management

  - Control and management information exchanged between the control and management entities with the eREC and the eRE. This information flow is given to the higher protocol layers and is considered as not being time critical.

- Synchronization:

  - Synchronization data used for frame and time alignment.

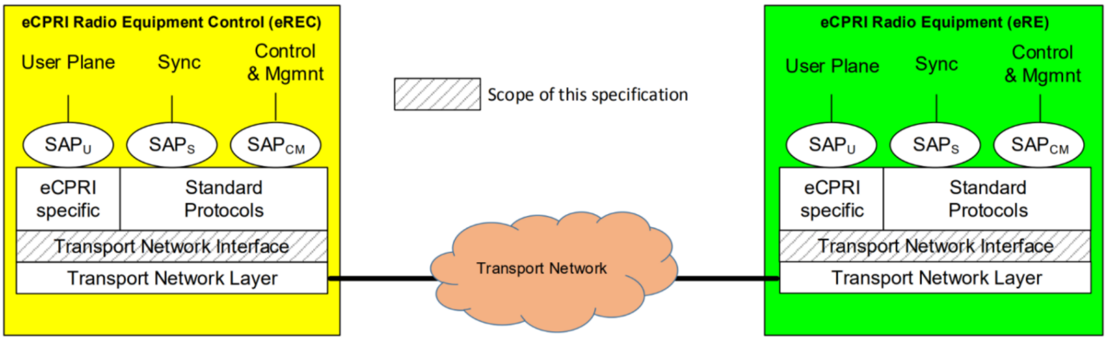


Figure 7-10 eCPRI architecture: System and Interface Definition

eCPRI defines a protocol for the transfer of user plane information between eREC and eRE via a packet based fronthaul transport network. For C&M and synchronization information flows, existing protocols and standards are referenced as proposals.

eCPRI does not constrain the use of specific network- and data link-layer protocols to form th network. Any type of network can be used for eCPRI, provided eCPRI requirements are fulfilled. “Requirements for the eCPRI Transport Network” aim to ensure that eCPRI system can use packet based transport network solutions and comply with the requirements associated with the more stringent radio technologies features in terms of timing and frequency accuracy, bandwidth capacity, latency, and packet loss etc.

The interface supports Ethernet-switched or IP-routed fronthaul networks. In case of eCPRI user plane over Ethernet directly, eCPRI messages shall be transmitted in standard Ethernet frames. The type field of the Ethernet frame shall contain the eCPRI Ethertype. In case of eCRPI user plane over IP, eCPRI messages shall be transmitted in UDP/IP packets.

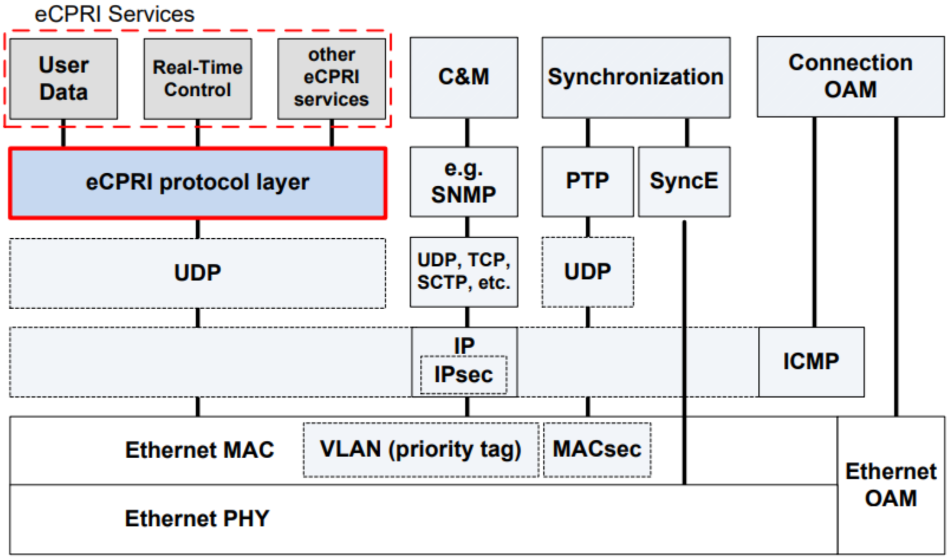


Figure 7-11 eCPRI protocol stack over IP/Ethernet

eCPRI enables flexible functional decomposition while limiting the complexity of the eRE. Split points located at the PHY level is one set of examples covered in the eCPRI specification.

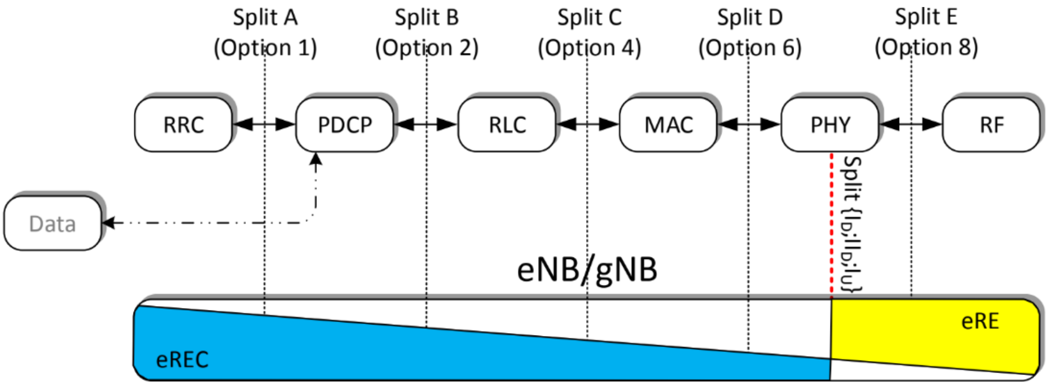


Figure 7-12 eCPRI node functional decomposition

The eCPRI specification focuses on three different reference splits, two splits in downlink and one split in uplink (I_D, II_D, I_U). Any combination of the different downlink/uplink splits is possible. The major difference between Split I_D and Split II_D is that the data in Split I_D is bit oriented and the data in Split II_D and Split I_U is IQ oriented.

Split points D, I_D, II_D, I_U and E are examples covered in the eCPRI specification. Split points Option 6, 7-1, 7-2, 7-3 and 8 are corresponding 3GPP split options and sub-options.

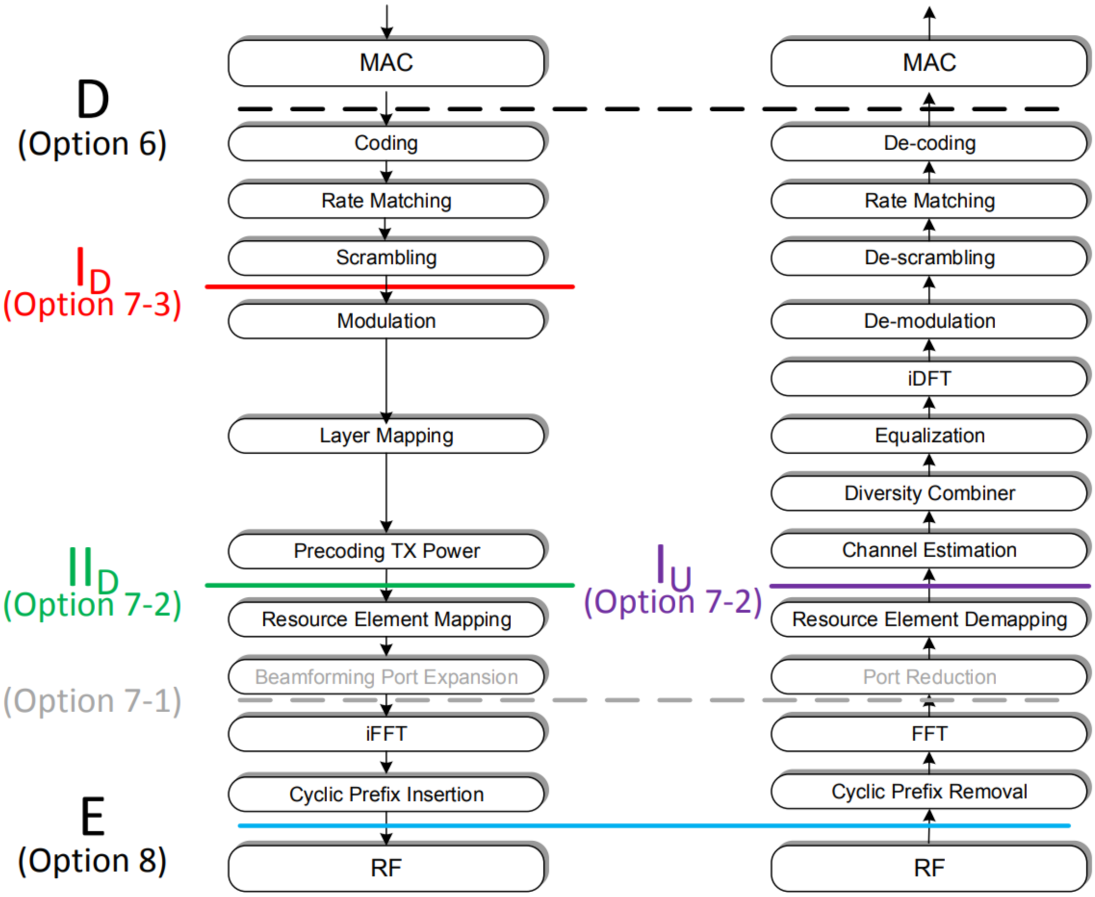


Figure 7-13 eCPRI functional decomposition

One of the major objectives of a new functional split between eREC and eRE compared to the classical CPRI functional split is to lower the bit rates on the fronthaul interface. When looking at the different processing stages performed in the PHY-layer in downlink direction, three processes will mostly increase the bit rate. These three processes are modulation, the port-expansion being done in combination with the beamforming process and the IFFT+cyclic-prefix-process (i.e. going from the frequency domain to the time domain). By moving the split upwards, the fronthaul bit rate will be lowered and vice versa. But conversely the bit rate for the user plane real-time control data will increase when moving the split point towards the MAC layer and vice versa.

7.6. Integrated Access and Backhaul

One of the potential technologies for future 5G cellular network deployment scenarios is wireless self-backhaul, which can enable flexible and very dense network deployment without the need for densifying the wired transport network accordingly. Compared to LTE, 5G NR can achieve much wider bandwidth and offer much higher throughput and network capacity through deployment of massive MIMO and multi-beam systems. It is therefore possible to develop and deploy integrated access and backhaul (IAB) links where relay nodes share the same radio resources with the macro donor access links to provide backhaul for other IAB nodes.(**错误!未找到引用源。**)

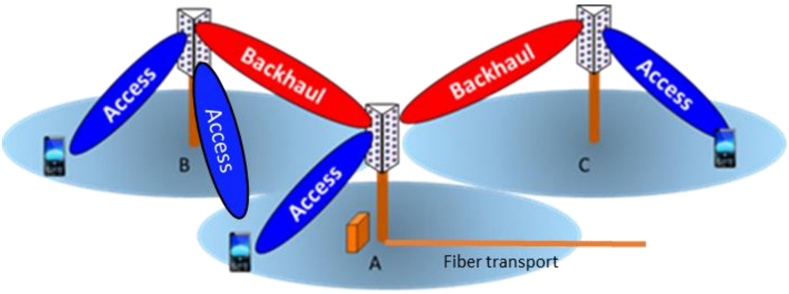


Figure 7-14 Integrated Access and Backhaul for 5G NR

It is envisioned that integrated access and backhaul can be used for both 5G NR deployments for both indoor and outdoor. Both in-band and out-of-band backhaul with respect to access link can be supported. In case of in-band backhaul, access and backhaul links will at least partially share the same frequency spectrum, so interference becomes the main concern. To prevent interference, there are two options for in-band backhaul solutions as shown in **错误!未找到引用源。**

1. The IAB nodes apply a half-duplex scheme, such as TDM (Time Division Multiplexing), FDM (Frequency Division Multiplexing) or SDM (Spatial Division Multiplexing), between access and backhaul links to avoid interference.
2. The IAB nodes apply a full-duplex scheme to allow simultaneous transmission and reception on both access and backhaul links, meanwhile add a self-interference cancellation module to mitigate the cross-link interference.

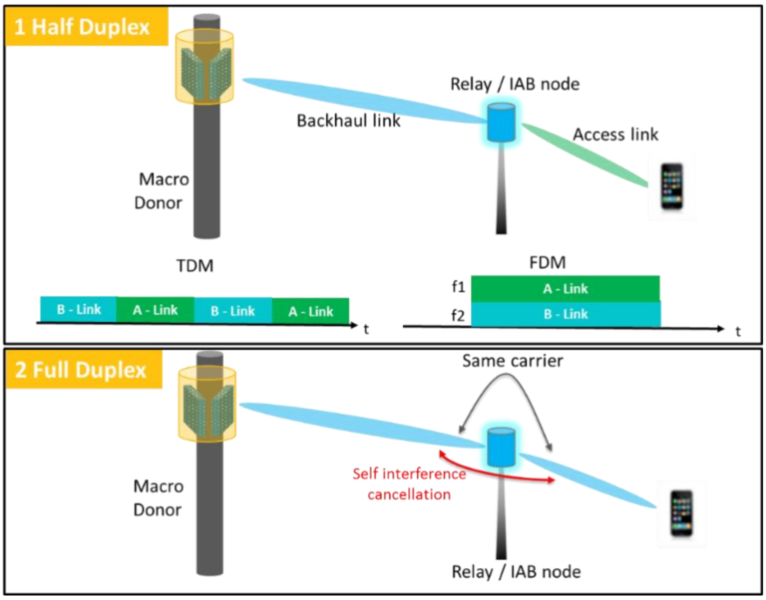


Figure 7-15 Two options for in-band backhaul solutions

Based on most recent 3GPP document of Dec. 2018(TR38.874 v1.0.0), In-band IAB subject to half-duplex constraint at the IAB shall be supported, but this requirement does not exclude full duplex solutions to be studied. Compared to half-duplex solutions, full duplex IAB node uses frequency spectrum much efficiently and creates much lower backhaul latency due to the fact that the IAB can transmit and receive simultaneously. On the other hand, interference mechanisms have to be implemented to make it work. In addition to self-interference cancellation, the interference between access and backhaul links can be further reduced through spatial separation and polarization isolation of transmit and receive antennas of the relay nodes and the small cell, as shown in 错误!未找到引用源。2.

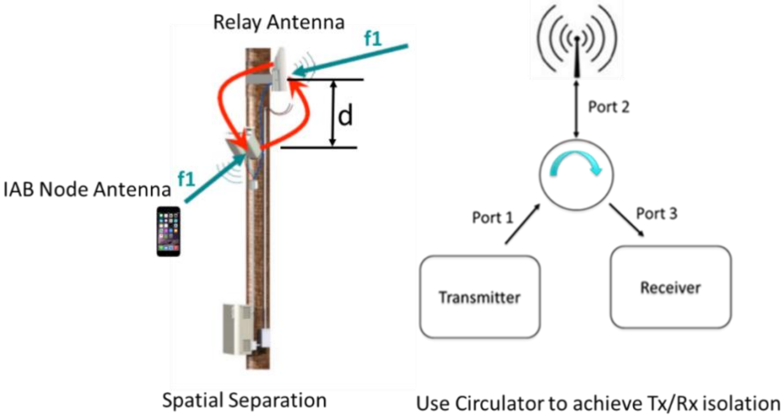


Figure 7-16 Antenna Separation to reduce interference

Massive MIMO system deployed on both donor gNB and IAB relay nodes make it possible to use beamforming algorithms for cross-link interference cancellation. Multiple beams can be formed to provide spatial separation of backhaul and access links and facilitate SDM of these links.

Multi-hop backhaul will be supported to provide range extension and redundant connectivity. Due to their limited range, this is especially beneficial for millimeter wave bands when they are used for backhaul links. Multi-hop also provides multiple options for backhaul routes. Autonomous adaptation on wireless self-backhaul network topologies shall be supported to minimize service disruptions and optimize backhaul capacity. Backhaul topology optimization algorithms shall be implemented to select the best route based on traffic load, signal strength, offered backhaul capacity and latency.

8. Solution

8.1. Cost Effective and Flexible Solution

The solution of C-RAN is anticipated as the evolution for cellular network topologies, where capacities are allocated and distributed among base stations through software means. The software-based nature of C-RAN is thought as a cost-effective approach to replace the traditional RAN architecture to enable new serviceability and simplified deployment approach.

SDN which decouples control plane and user plane and NFV which makes a decoupling of software and hardware together make the architecture extremely flexible. Radio Functions can be realized flexibly and re-usably with software defined method, a software implementation of network functions on top of GPPs with no or little dependency on a dedicated hardware. Thus, the flexibility offered by SDN method improve service life-cycle and across-platform portability at the cost of lower power and computational efficiency. Using NFV extends this flexibility through virtualization of the execution environment, execution of network functions on top of vitualized (and shared) computing, storage, and networking resources. Consequently, radio functions become a general purpose application that operates on top of a virtualized environment and interacts with physical resources either directly or through a full or partial hardware emulation layer. The resulted virtualized software radio functions can be delivered as a service and managed through a cloud controller. This changes the mobile networks towards cheap, flexible and easy to manage software platforms.

The architecture of C-RAN also leverages the use of white-box and GPP hardware servers that can work flexibly with software-defined radio network functions which can be virtualized for easier management by adopting NFV (Network Function Virtualization) and cloud computing orchestration. The white-box hardware performs general functions and signal conversions. No more proprietary ASIC or DSP is required when new generation of hardware interface is released. At the same time, virtualized components ensure real-time performance by dynamically allocating hardware resource of the white-box server. The could RAN offers higher flexibility and saves the operating costs for cross-compatibility and hardware life-cycle.

8.2. xhaul Solution

When building the mobile xhaul transport network for 5G, the preferable technique chosen depends on the availability and cost of optical fibre connections. If fibre is highly available or may be easily deployed, point-to-point fibre is a favourable choice imposing no capacity limits and lowest delay. If fibre is available but costly, sharing the fibre, e.g. using WDM-PON (Wavelength Division Multiplex - Passive Optical Network) and Time Sensitive Network (TSN) based Ethernet switching solutions may be a more cost effective alternative. At Some places fibre is not available and too expensive to deploy. In such environments, microwave or wireless optical links may then be the solution.

In a location like a stadium or enterprise building, a fiber/Ethernet based fronthaul would enable the benefit of C-RAN for small cell deployment, leading to better performance achieved by higher coordination among small cells and lower cost of utilizing a shared baseband processing pool. In this case, certain fronthaul compression mechanism will have to be used to reduce the bandwidth requirements while meeting the latency target.

For radio interface data transport, CRPI is mostly used in LTE while eCPRI is designed for 5G NR. As previous state, eCPRI is a packet based technology with encapsulation which can support larger data volume than CPRI at relative lower cost. However, no matter CPRI or eCPRI, they both can be supported by the fronthaul transport solutions mentioned above.

In scenarios without fibre based midhaul/backhaul, wireless backhaul and Integrated Access and Backhaul (IAB) could be a better solution, as discussed in the previous section.

8.3. Coverage and Capacity Solution

Indoor 5G target networks require at least one frequency band to achieve continuous coverage.

The radio propagation loss and penetration loss of Sub-6GHz bands are much lower than those of mmWave, so Sub-6 GHz bands can achieve continuous indoor 5G network coverage at a relatively low network construction cost. Meanwhile, AAU/RRU with high density are also required to achieve continuous coverage. mmWave features wide spectrum but weak coverage capability. Indoor areas where the Sub-6 GHz spectrum resources cannot meet service requirements, can overlay mmWave onto these networks to meet the requirements of ultra-large capability.

Considering spectrum resources, radio propagation characteristics, and network construction costs, Sub-6 GHz is used to provide continuous indoor coverage for 5G basic coverage and capacity layer. While, in hotspot areas mmWave spectrum can be used for traffic absorption and capacity extension.

In addition, MIMO technology which supports multi-layer transmission over the air interface and provides uplink and downlink diversity gains, greatly increases cell capacity and improves network edge rates and user experience. In addition, MIMO technologies should be used to cancel co-channel interference among neighboring cells (between macro and small cells, among small cells)

Antenna form is one of the main limiting factors in applying MIMO technologies in small cells. A realistic antenna form factor for commercialized small cell ranges from 150mm – 250mm. To achieve good MIMO gain, it is generally required that the separation of the antenna elements is of the order of half wavelength. In case of 2.5 GHz carrier frequency, the wave length is 120mm, a 4T4R unit will result in a form factor of 180 mm. An 8T8R unit will drive the antenna unit size to 420mm, which is way too big to be installed on a light pole or on the ceiling of a room. The antenna form factor might be a bit smaller for 3.5 GHz due to shorter wavelength, but an 8T8R unit would be ~300 mm, - way too big for commercial deployment. Considering the antenna form factor volume, technical complexity and terminal specification, massive MIMO might not be a suitable solution for indoor scenario of small cells at sub-6 GHz bands. 4T4R is a more appropriate choice to provide system throughput and coverage gains, which meets the requirements of ultra-high bandwidth, especially ultra-high traffic density and the cell edge rate of DL 100Mbps.

Interference cancellation between macro cell and small cells as well as among small cells is one of the biggest challenges in successful small cell deployment. When macro-cells and small cells operate at the same carrier frequency, small cells will suffer strong DL interference from neighbouring cells (either macro or small cells). If the handover boundary between cells is based on the received signal strength at UE, many UEs close to the cell boundaries of small cells are actually located in the service area of a macro cell. This will cause very strong uplink interference. It is therefore very important to implement effective interference cancellation mechanisms in order to realize the full benefit of small cells.

8.4. Elastic Capacity Solution

Indoor 5G network capacity planning concentrates on the traffic model, crowd density, and target coverage area. The traffic model is considered the most important factor. During initial network construction, the forecast for 5G traffic models and capacity requirements in different scenarios can be based on 4G networks and the development trend of 5G services.

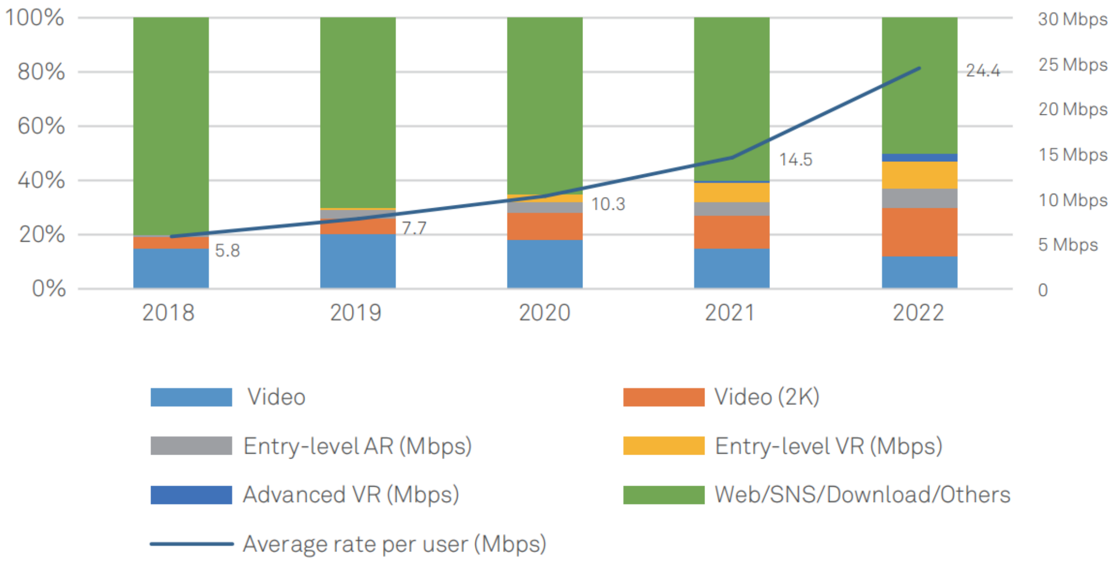


Figure 8-1 5G traffic model forecast

As shown in the preceding figure, high bandwidth services such as VR services and HD videos are rapidly increasing. As said in the previous context, the traffic burst varies as time or location changes. Therefore, elastic and flexible capacity design is a serious issue that must be considered during indoor 5G network capacity planning.

- Elastic capacity design shall predict 5G capacity requirements based on scenario characteristics, 5G service characteristics, historical network development data, and 5G user development plan.
- Network architecture design must comply with the flexible capacity expansion principle. Equipment space and transmission lines must be reserved for several times of elastic capacity expansion.
- Network layout must ensure that no obvious interference exists between cells after capacity

expansion.

- When designing an elastic capacity network, 5G traffic burst areas, such as news centers and assembly sites, must be considered.

Using CU/DU or CU/DU/RRU split architecture, capacity can be flexibly configured as required due to the headend-level cell splitting and merging capability.

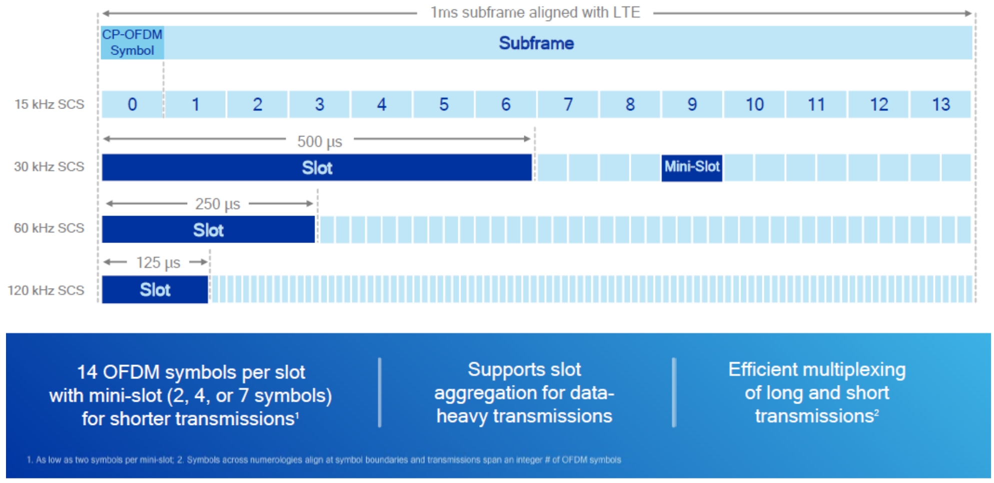
8.5. Low-latency Solution

To meet URLLC requirements for reducing latency requires reducing radio signal transmission time and reducing time needed for retransmissions.

1) Reducing radio signal transmission time: the longer the slot length, which is the unit of signal transmission, the longer the wireless signal transmission time will be. So reducing the slot length is desirable. 5G NR new radio frame structure have introduced multiple different Orthogonal Frequency Division Multiplexing (OFDM) subcarrier spacing (15, 30, 60 and 120kHz). Using wide OFDM subcarrier spacing provides wider bandwidth per subcarrier, so that the same amount of information can be transmitted in a shorter time. This enables the transmission time for the radio signal to be shortened to achieve low latency. Moreover, mini-slot has been specified in the NR specification occupying 2, 4 or 7 symbols, which can be used as the minimum scheduling unit further reduce latency.

2) UL grant free transmission can also be used to reduce latency. UE will perform K grant-free consecutive transmission of the same transport lock on the pre-configured grant-free resources. UE will keep the repetition of the transmissions until having received an ACK or having reached the K transmissions.

3) DL preemption indication based DL eMBB and URLLC multiplexing can be implemented to ensure dynamic resource sharing between URLLC and eMBB. Through pre-empting eMBB data in some slots and transmitting URLLC traffic instead when needed, the ultra-low latency service will be guaranteed.



14 OFDM symbols per slot with mini-slot (2, 4, or 7 symbols) for shorter transmissions[1]

Supports slot aggregation for data-heavy transmissions

Efficient multiplexing of long and short transmissions[2]

1. As low as two symbols per mini-slot; 2. Symbols across numerologies align at symbol boundaries and transmissions span an integer # of OFDM symbols

Figure 8-2 Scalable 5G NR slot duration for diverse latency/QoS

4) Reducing time needed for retransmissions: Successful packet reception rates can be increased through signal retransmission, but the retransmission procedure require signaling on the uplink and downlink, which introduces delay. Thus it is important to reduce the time needed for retransmission in order to achieve both high reliability and low latency. One of the solution is TCP performance enhancement through prioritizing TCP retransmission. In case of bidirectional TCP transmissions, it is possible that TCP ACK/NACKs for DL traffic are delayed by the UL TCP data, which will slow down the retransmission and degrade the DL TCP performance. On the other hand, the high DL data rate will generate excessive TCP ACKs on the UL, which will block the actual UL data traffic. Some enhancements are needed to enhance TCP performance. One of them is prioritizing TCP ACK transmission. The PDCP inspects the contents of the PDCP SDU, and prioritizes the TCP ACK by changing the transmission order of TCP ACK packets earlier than TCP packets so that TCP ACK packets can be transmitted as early as possible.

8.6. Reliability Solution

Reliability can be improved through a secure and always-available network with built-in redundancy.

Indoor 5G network coverage redundancy must be considered and multiple headends must be deployed in areas requiring high reliability. If one of the headends malfunctions, the adjacent headends can provide complementary signal coverage.

Also indoor 5G network capacity redundancy must be taken into consideration. If a single headend cannot provide sufficient capacity, the adjacent headends can be scheduled to provide extra capacity on demand.

Moreover, network reliability also depends on the network topology during network design. Redundancy backup is a key factor for critical connection and nodes. When a temporary fault occurs, other backup nodes complement the faulty one to maintain proper functioning.

To ensure higher reliability of the system, some new software features can also be implemented. For example, PDCCH repetition or larger AL (Aggregation Length) for control channel can be used for purpose of improving reliability of control channel and reducing the BLER. In addition, the existing MCS/CQI matching table (for BLER of 10^(-1)) shall be updated into a new MCS/CQI matching table shall be supported in order to support 256 QAM with the target BLER of 10^(-5) based on 3GPP Release 15 standard. Of course, both macro and small cells will have to implement the feature at the same time.

9. Deployment Options

There are multiple deployment options for indoor scenarios, such as Integrated Base Station and Distributed Indoor System, e.g. DAS, Extended pRRU architecture and Distributed pRRU architecture.

Integrated base station can be implemented to provide deep coverage for small area or to support supplementary capacity for hotspots.

Distributed indoor system can be categorized into three classes: DAS, extended pRRU architecture and distributed pRRU architecture according to the capacity provided.

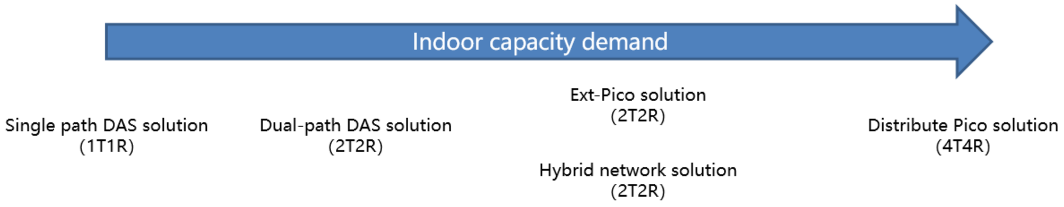


Figure 9-1 Indoor deployment solution related to capacity demand

Table 9-1 Distributed indoor system deployment options

	Re-use existing facilities	Deploy new facilities
Passive	Option 1: add a high power 5G NR RRU at the	Deploy high power 5G NR RRU

DAS solutions	place of signal source Option 2: add a high power 5G NR RRU and a Smart GW at the place of signal source, replace legacy DAS antenna with Smart DAS antenna (battery powered)	as the signal source, DC power transparent passive facilities and Smart DAS antenna (DC powered)
Active pRRU solutions	Option 1: add 5G single-mode active pRRU. Option 2: replace legacy 4G pRRU with 5G multi-mode active pRRU.	Deploy 5G multi-mode active pRRU

Besides the indoor deployment options mentioned in the table can be deployed based on the requirement, a hybrid solution which mixes Distributed pRRU system with DAS can also be used, e.g. distributed pRRU (without antenna) as signal source is deployed with DC power transparent passive facilities and Smart DAS antenna (DC powered)

NOTE: pRRU mentioned in this Whitepaper consists of both RRU part and antenna part, unless otherwise stated.

9.1. DAS Deployment with Smart DAS Antenna Solution

With innovative integration of DC power transparent passive facilities and Smart DAS antennas, the total CAPEX and OPEX of 5G indoor deployment could be saved and other extra services can also be provided such as indoor positioning and smart O&M with the help of like BLE .

In this deployment option, high power 5G NR RRU is deployed as signal source with either existing passive facilities reused with battery powered Smart antenna or DC power transparent passive facilities with DC powered Smart antenna.

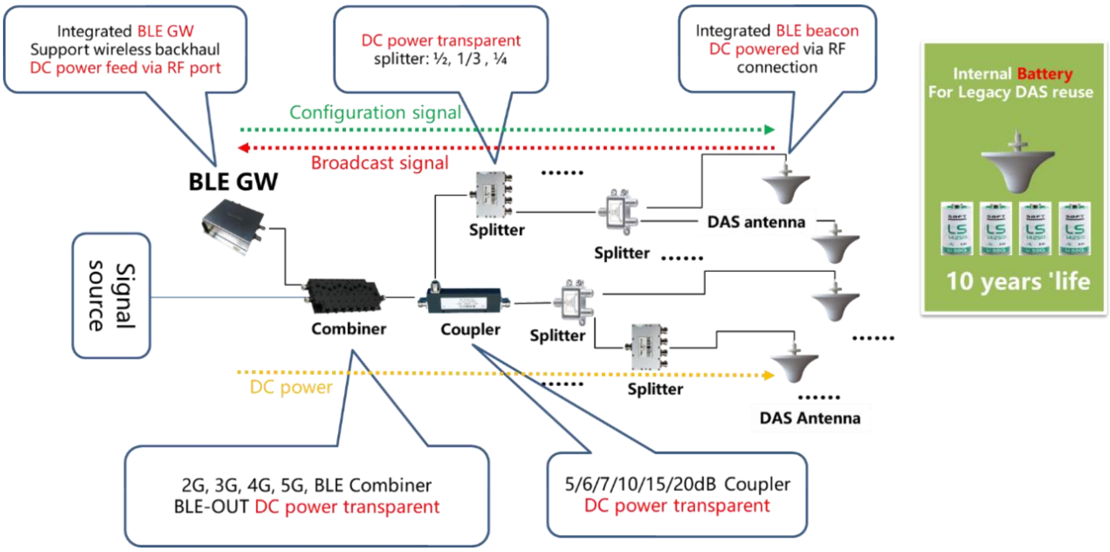


Figure 9-2 Example of Smart DAS indoor deployment with the help of BLE

9.2. Distributed Indoor Deployment Solution

Extended pRRU architecture and distributed pRRU architecture are two prime architectures for Distributed Indoor solution deployment nowadays. These two architectures have the same network architecture which consists of three parts: Base-Band Unit (BBU), Extender Unit (EU) and Remote Unit (RU), but with different capabilities to satisfy different scenarios.

The Extended pRRU architecture inferior to the distributed one lies with the lower configuration for the process capability of the Base-band Unit which cannot support cells splitting to enlarge the whole system’s capacity. Irrespective of this difference, these two solutions are almost the same.

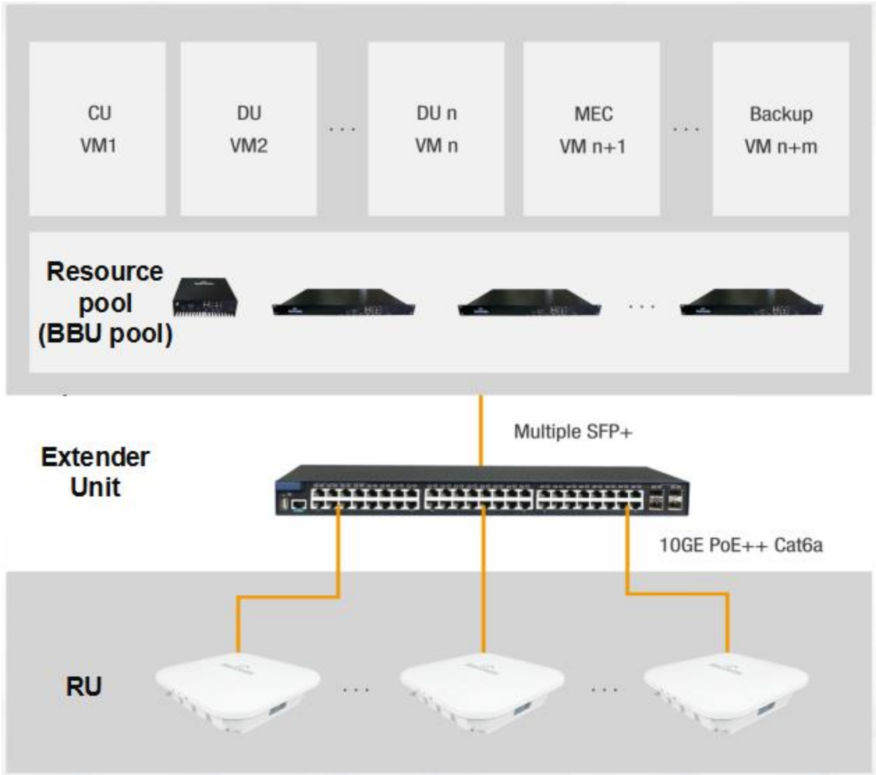


Figure 9-3 Example of Distributed Indoor Deployment Solution

1. BBU

BBU can be deployed based on GPP with SDN/NFV technology virtualizing BBU functions which will bring in pooling effect to achieve most efficiency of resources usage. The specific functions to be realized is the similar to the traditional BBU functions. BBU can also be split into CU and DUs which are deployed distributively.

2. Extender Unit

The connection between extender unit and RU can be built based on twisted pair cable (e.g. Cat6a), optical fibre or hybrid cable depending on the distance between these two units. For example, distance below 100m, Cat6a is preferred, longer than 200m, optical fibre will get more consideration, and other distance, hybrid cable is recommended.

Additionally, because of the higher speed need in 5G, at least 10Gpbs+ per cell optical module shall be supported at the interface.

3. Remote Unite (pRRU)

pRRU can be implemented based on white-box hardware to take full advantage of the economies of scale offered by an open computing platform approach. This remote unit can be co-located with 4G RRU with more power to provide the same coverage. Or deploying 5G pRRU densely providing relative smaller coverage needs less power per pRRU.

9.3. Hybrid Solution with Smart DAS Antenna

In this Hybrid network solution, Extended-Pico or Distributed Pico 5G pRRU (without integrated antenna) with 4 external RF ports is used as the signal source, additionally, the BLE GW can be deployed at the same position with pRRU providing RF system monitoring function and power feed to smart DAS antennas. pRRU could be 5G single mode or multi modes to fulfill different access network needs. Single polarized passive DAS antennas could form single path 1+1+1+1 coverage configuration and dual polarized passive DAS antennas could form dual path 2+2 coverage configuration, so that 2x2 or 4x4 MIMO could be realized at the overlapping areas.

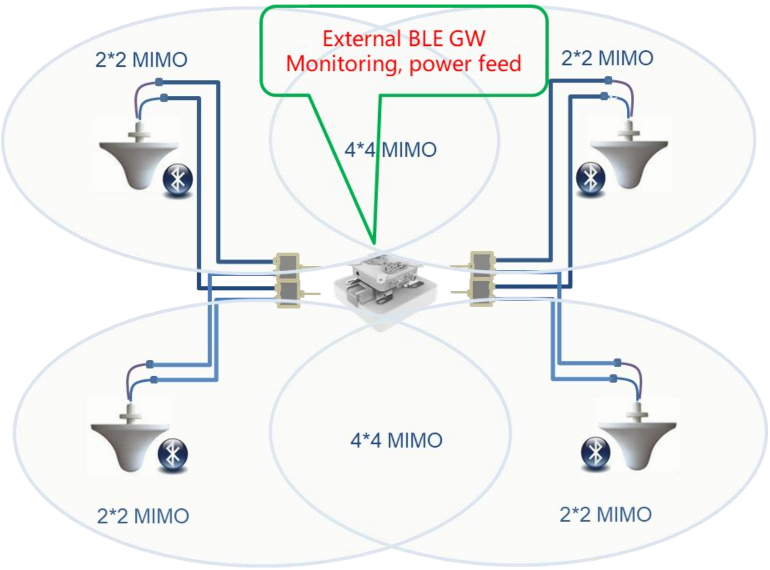


Figure 9-4 Example of hybrid solution deployment

10. Summary

Indoor 5G services such as ultra HD video, VR and massive intelligence sensor interconnection require capabilities such as full coverage, high bandwidth, low latency, high reliability, scalability and flexibility etc,. Also, indoor 5G network faces challenges and requirements of high-band networking, elastic capacity, network low latency and redundancy, scalability and flexibility with cost-effective deployment.

Thus, the cost-effective C-RAN solution is expected as the evolution for cellular network topology to settle the challenges and meet the requirements. C-RAN architecture leverages software defined method, a software implementation of network functions on top of white-box and general purpose hardware servers with no or little dependency on a dedicated hardware, and virtualization of execution environment, execution of network functions on top of vitualized and shared computing, storage, and networking resources which will provides scalability and flexibility to the deployment.

C-RAN can consist of CU, DU and RU two or three parts. RU can be implemented based on programmable white-box hardware while DU and CU can be built on GPP and FPGA hardware with higher layer radio functions which are virtualized as VNFs. Transmission between CU, DU, and RU, so called xhaul, fibre transmission is the ideal solution, but in the fibre-less scenario, we can use wireless solution as a substitution, e.g. IAB for midhaul and backhaul. eCPRI can be used in the Fronthaul between RU and DU which provides compression mechanism to reduce the bandwidth required while meet the latency target to some extent.

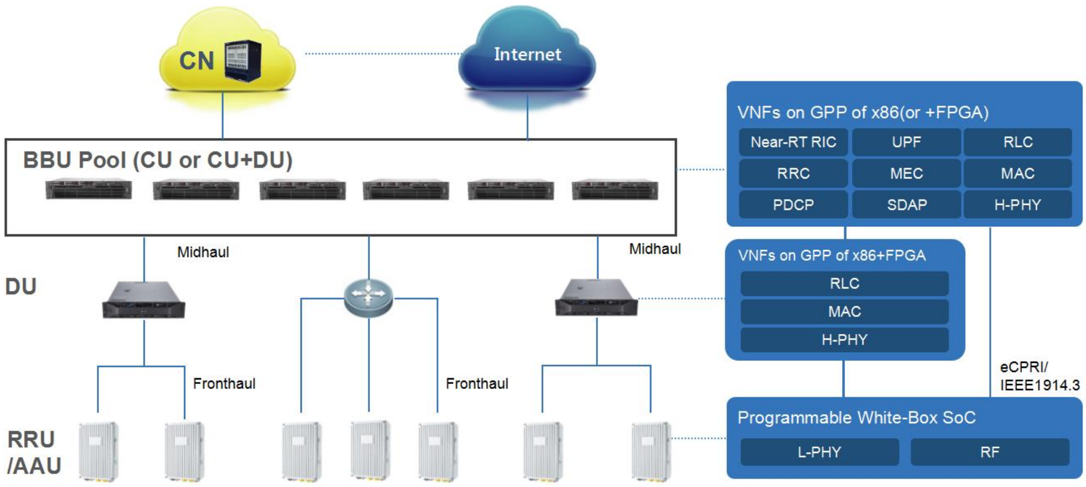


Figure 10-1 Example of C-RAN architecture

Based on C-RAN architecture, cell splitting and merging technology can be used to provide elastic cell capacity to settle tidal phenomenon. Also we recommend that for high-band indoor networking, frequency in sub-6GHz can be used to provide coverage, and mmwave can be used in the hotspot as a complement of sub-6GHz for ultra-high capacity requirement with the help of recommended 4T4R MIMO and interface cancellation technologies.