Mobile Virtual Network Operator (MVNO)

Why MVNOs

A MVNO (*Mobile Virtual Network Operator*) offers wireless communications services by leasing network capacity and other services from a traditional MNO (*Mobile Network Operator*), rather than owning the infrastructure itself

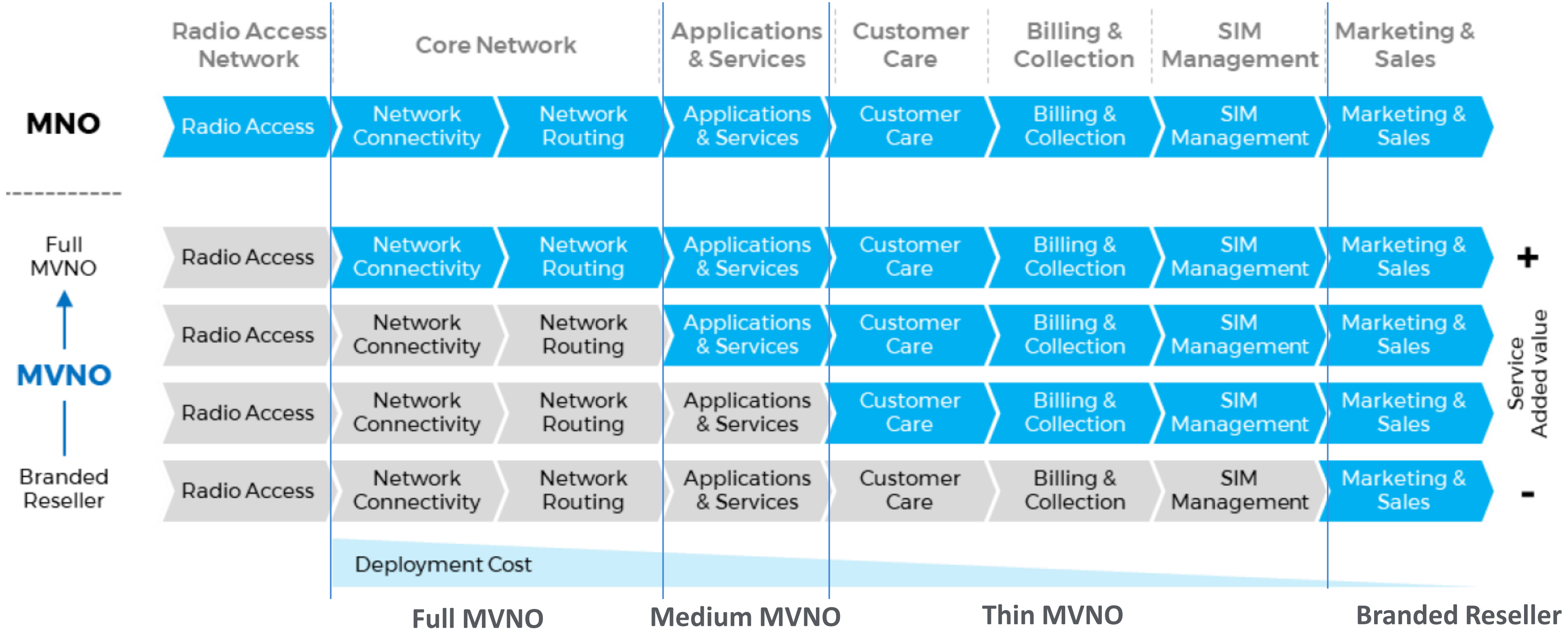
Business model which allows for competitive pricing and specialized services for consumers and businesses alike

5G provides:

1. New in business opportunities (mMTC, URLL and NPN)
2. New, more spectrum, thus more resources for all
3. Slicing exploitation to create virtual networks
4. More modular architecture
5. eSIM (not directly related to 5G!), avoiding the need for physical cards

Mobile Virtual Network Operator (MVNO)

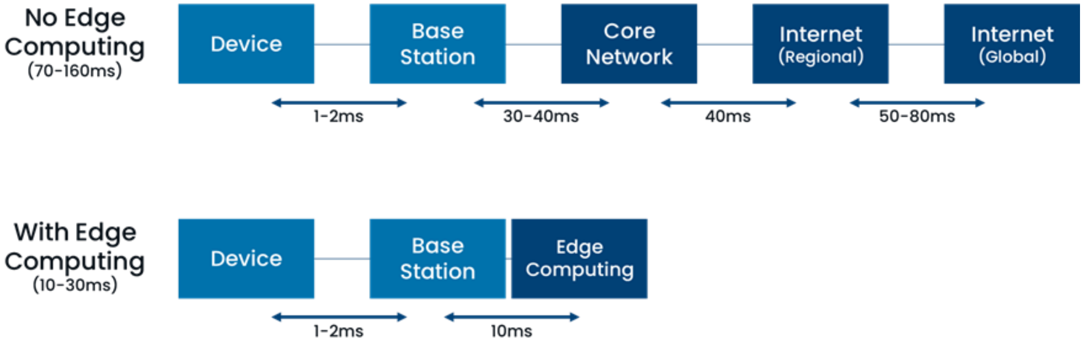
Types of MVNO:

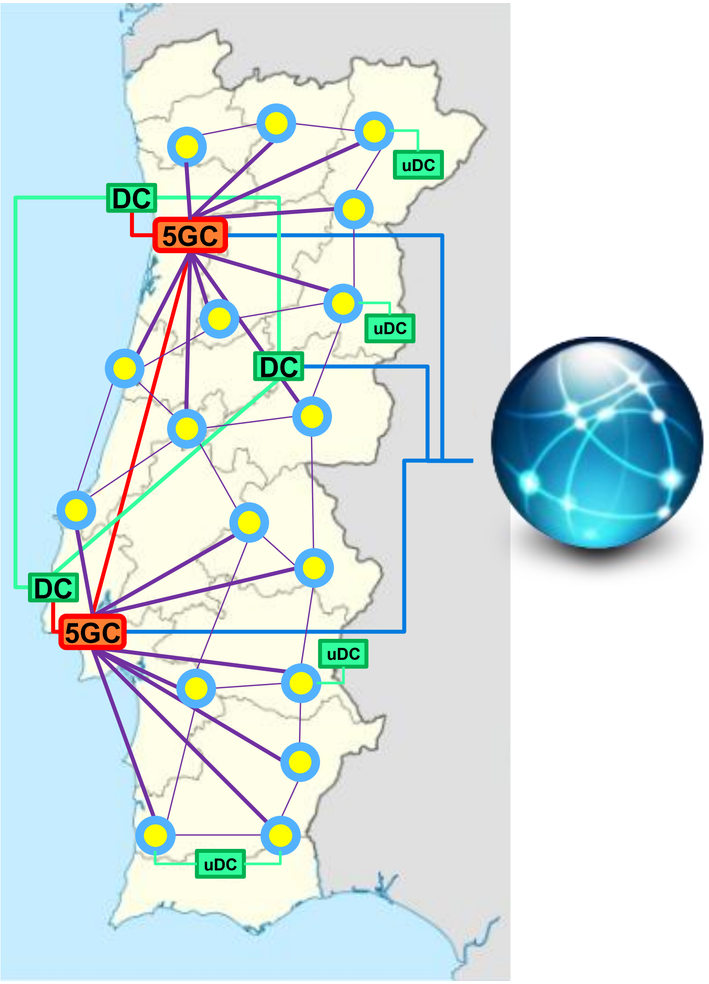


Edge computing

Distributed cloud: Edge Computing and 5G

- Services (operator and 3rd-party Edge Applications) requiring low latency are not compatible with central or cloud computing
- The overall latency shall be as low as the one imposed by the radio interface (5G-NR)
- Solution
  - Distributed, small data centres (NFV powered), placed close to the network edge

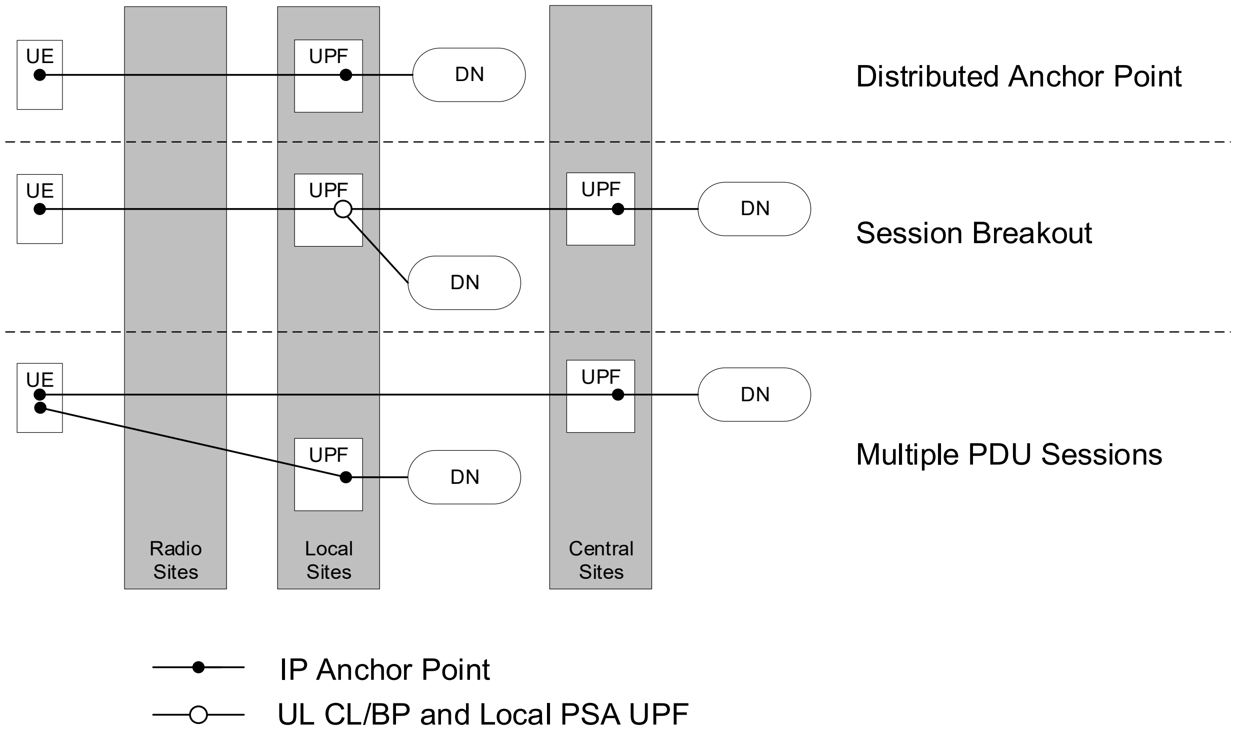




- Also allowing processing offloading from UEs
- 5G provides native support for (*Multi-access) Edge Computing* (MEC)

Distributed cloud: Edge Computing and 5G

Edge Computing enables operator and 3rd party services to be hosted close to the UE's access point of attachment, so as to achieve an efficient service delivery through the reduced end-to-end latency and load on the transport network (3GPP 23.548)



UL-CL/BP*: Uplink Classifier/Branching Point*

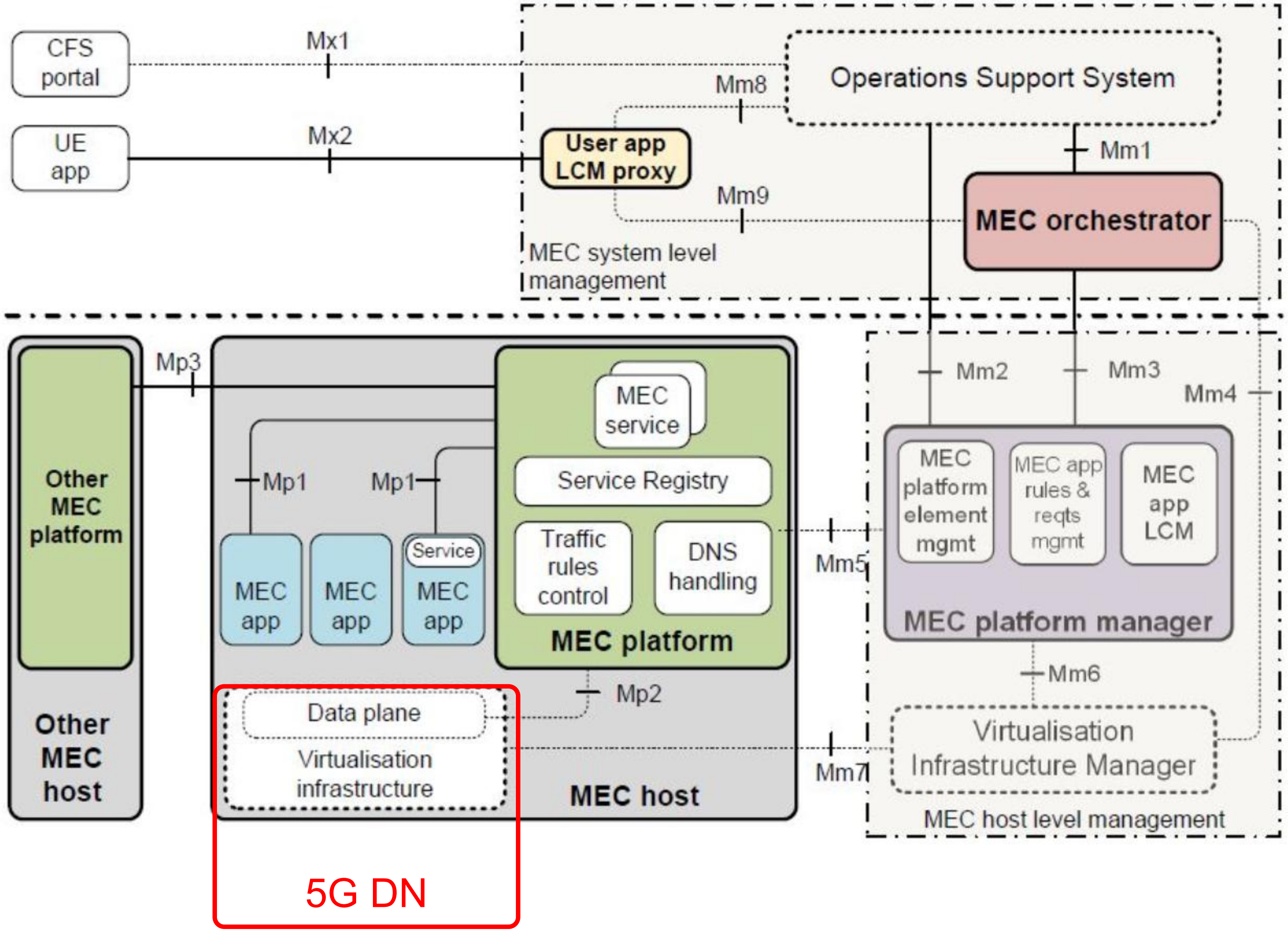
- Function within the *User Plane Function* (UPF) that handles sophisticated traffic steering for uplink data
- Key enabler for 5G edge computing

UL-CL: examines uplink traffic and directs it to the correct UPF based on its content or other attributes

BP: acts as a branching point for traffic, allowing a single flow to be split and sent to multiple destinations

PSA UPF: PDU Session Achor Point UPF

ETSI *Multi-Access Edge Computing* (MEC)



Provides an open environment allowing the efficient and seamless integration of applications from vendors, service providers, and third-parties across multi-vendor *Multi-access Edge Computing* (MEC) platforms

Services:
IMS and VoNR

IP Multimedia Subsystem (IMS)

Standards-based architectural framework for delivering multimedia communications services such as voice, video and text messaging over IP networks

Originally created by 3GPP to standardize the implementation of these services on next-generation mobile networks, with operator's guarantee of service – the customer pays the service not the traffic

NOT *Over-the-Top* (OTT) but integrated with the network, guaranteeing service delivery

Application Layer

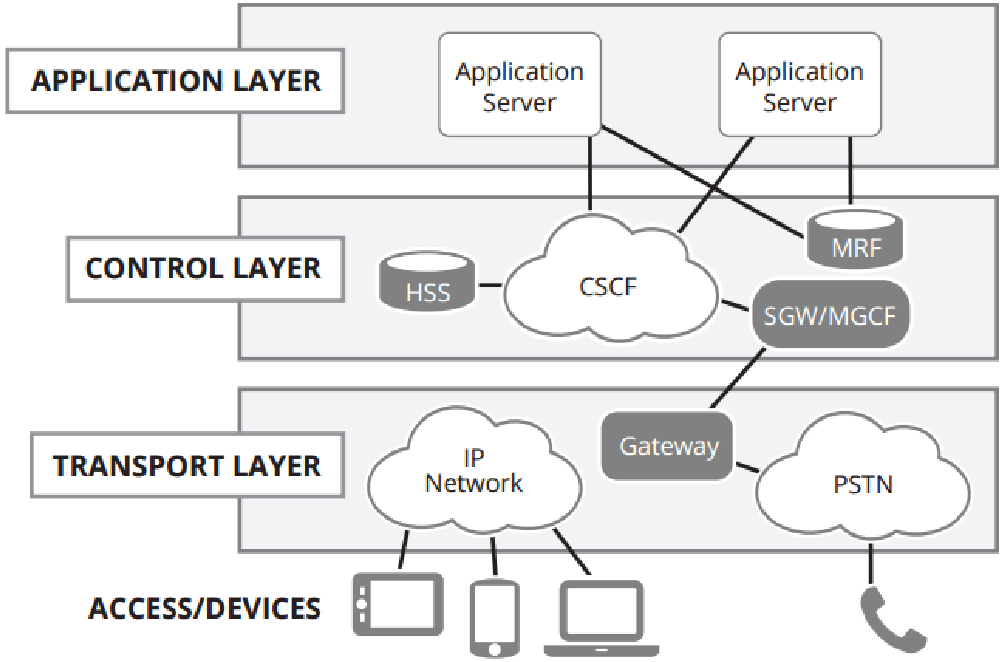
- **Application Servers**: services logic (e.g. telephone call control, call continuity, conference calling, and supplementary service configuration)

Control Layer

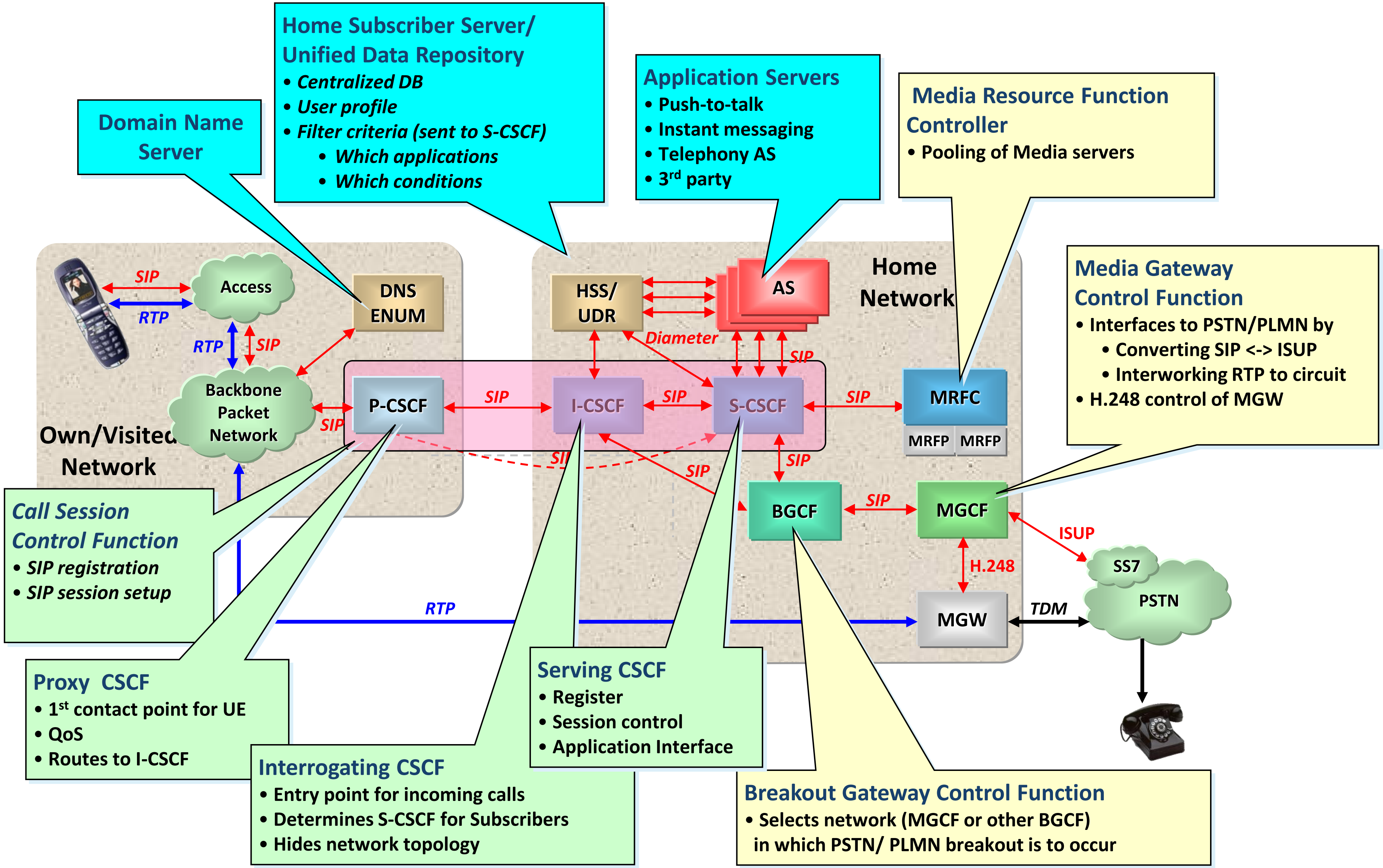
- ***Call Session Control Function* (CSCF)**: core of the architecture, sessions control between endpoints (referred to *User Agents* in the IMS specifications) and applications; crucial in ensuring seamless and secure communication services
- ***Home Subscriber Server* (HSS)**: the master database that maintains all user profile information used to authenticate and authorize subscribers

Transport Layer

- ***Signaling Gateway* (SGW) and *Media Gateway Control Function* (MGCF)**: respectively provide SS7 and MGCP interoperability with the PSTN
- ***Media Resource Functions* (MRF)**: provides media-related functions such as the playing of tones and digital announcements



IP Multimedia Subsystem (IMS)



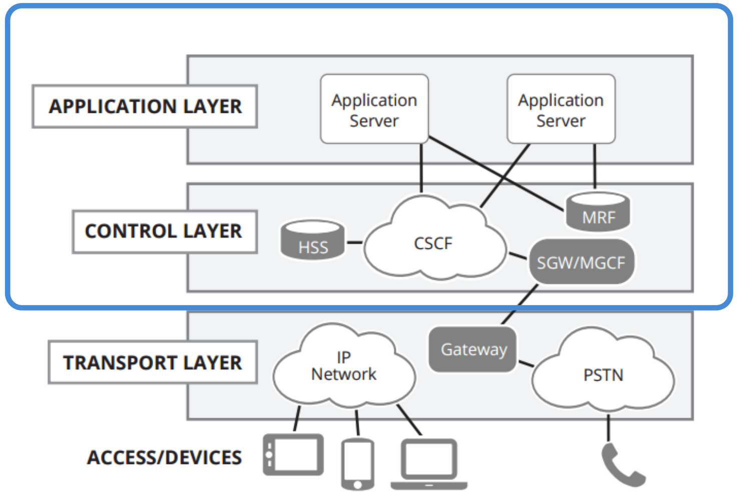
IP Multimedia Subsystem (IMS)

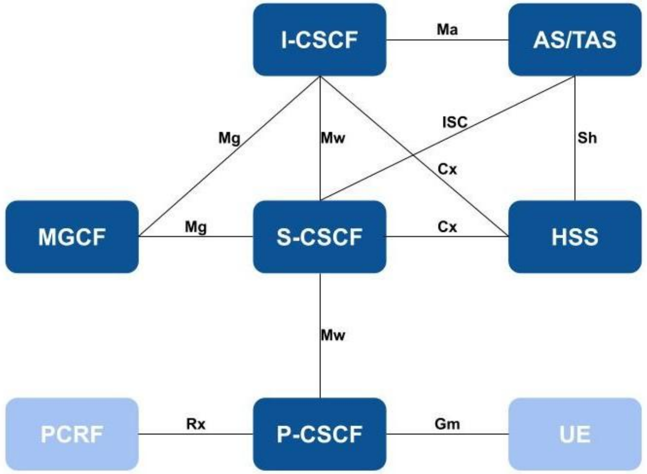
- **S-CSCF (*Serving Call Session Control Function*)**
  - The central node of the IMS network (only one logical entity in the Home domain)
  - Is a SIP Registrar
  - Handles session control and ensures proper SIP signaling for services
  - Retrieves subscriber data from the HSS/UDR and enforces service-specific rules

- **P-CSCF (*Proxy Call Session Control Function*)**
  - The first point of contact for a User Equipment (UE) in the IMS network; several distributed in the network may exist
  - Responsible for forwarding SIP signaling between the UE and other IMS entities
  - Located in the visited network
  - Terminals (UA) must know this proxy (e.g. via DHCP)
  - Performs security functions like encryption/decryption and compression/decompression
  - Interacts with the PCRF for QoS and policy enforcement

- **I-CSCF (*Interrogating Call Session Control Function*)**
  - Acts as an entry point into an operator's IMS network from external networks; one or more may exist per domain
  - Responsible for querying the HSS or 5G-UDM via the Cx interface to determine the user's location and routing SIP requests to the appropriate S-CSCF

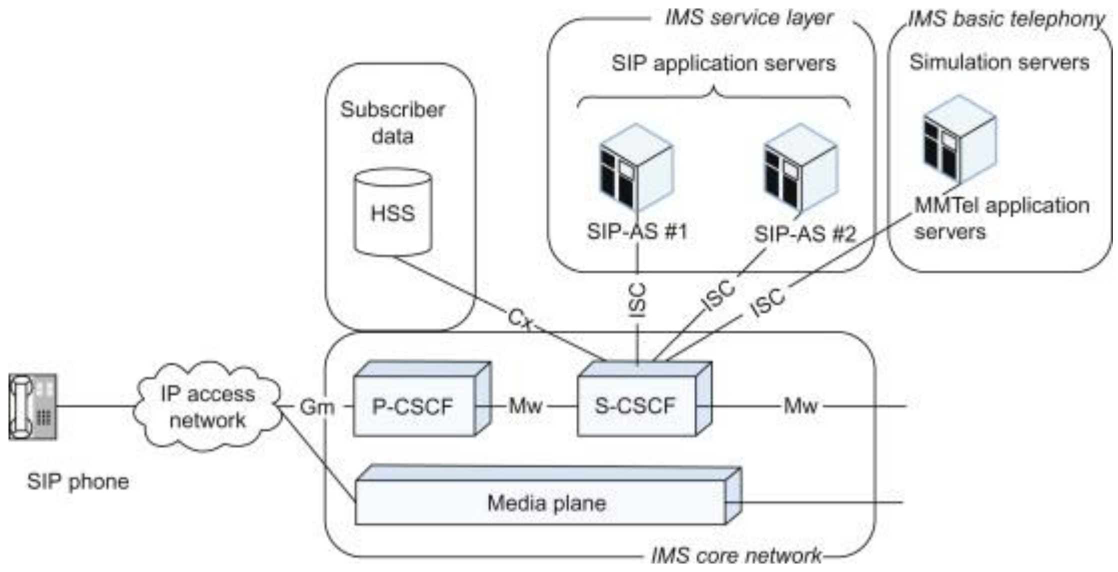
- **HSS (*Home Subscriber Server*) or 5G-UDM (*Unified Data Management*)**
  - A centralized database for storing subscriber profiles, authentication credentials, and service-related information
  - Interacts with the S-CSCF to provide user data for call/session handling





Services in IMS

- IMS is an advanced infrastructure enabling services
  - AS run the logic that define the provided services (VoLTE, VoNR, MMTel, IPCentrex, etc)
- Services are in the end points or peers (calls, etc.), not in the IMS
  - SIP *User Agents* (UA)
  - *Application Servers* (AS)
  - *Media Gateways* (MG) and *Border Gateways* (BG)
- AS are the key part to endow IMS with services
  - AS can behave as SIP proxy or as a SIP UA (terminal)
- AS interact, using SIP, with the S-CSCF (which controls user's SIP session)



SIP Protocol

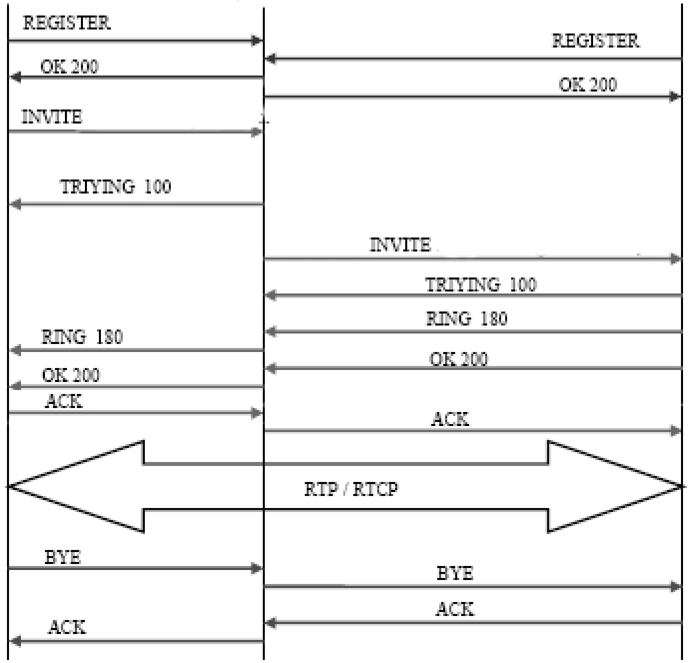
- SIP (*Session Initiation Protocol*): main signaling protocol for IMS, adopted to manage end-to-end multimedia sessions like voice and video calls
  - SIP: establishment, modification, and termination of those sessions
  - IMS: framework for service control, user management, and session management across different access networks
- Defined in IETF RFC 3261
  - "… an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include Internet telephone calls, multimedia distribution, and multimedia conferences."
- In IMS, SIP is extended to include extra functionality (e.g. 3GPP TS 23.228)

SIP Protocol

SIP Request messages

These are commands sent by a client to a server to initiate an action

- **INVITE**: Used to start a session
- **ACK**: Used to confirm the final response to an INVITE
- **BYE**: Used to terminate an established session
- **REGISTER**: Used to register the UA in core
- **CANCEL**: Used to cancel a pending transaction
- **OPTIONS**: Used to query a server for its capabilities



SIP Response messages

These are the replies from a server to a request, similar to HTTP responses

They are numerically coded to indicate the status of the request

- **1xx (Provisional)**: Indicates the request is being processed (e.g. *100 Trying*)
- **2xx (Successful)**: Indicates the request was successful (e.g. *200 OK*)
- **3xx (Redirection)**: Indicates the recipient must take further action to fulfill the request
- **4xx (Client Failure)**: Indicates an error on the client's part (e.g. *404 Not Found*)
- **5xx (Server Failure)**: Indicates an error on the server's part

IMS identifiers

IMPI (*IP Multimedia Private User Identity*) (as the SUPI at 5G level)

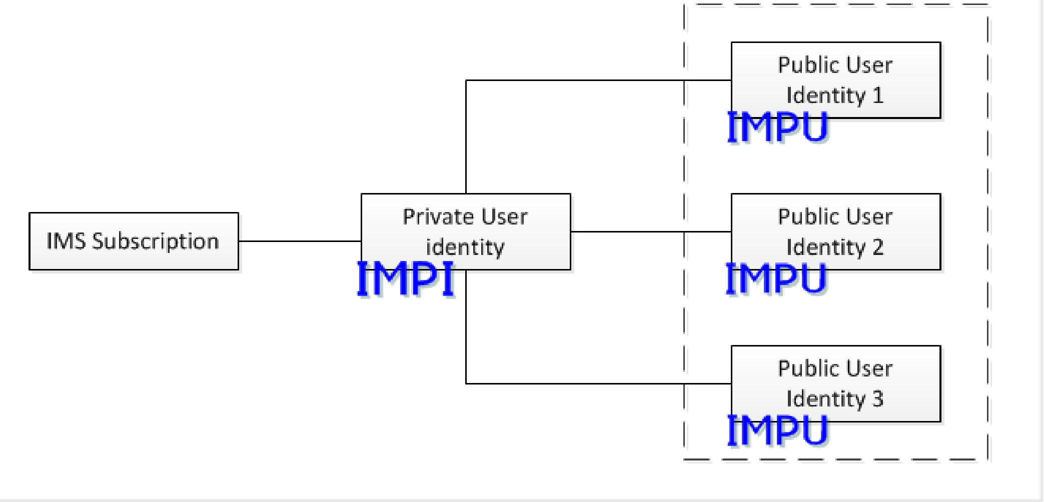
A private, permanently allocated global identity assigned by the home operator to the subscriber

- **Usage**: Used for authentication, registration, and administrative purposes within the network
- **Format**: A *Network Access Identifier* (NAI), such as user.name@domain
- **Storage**: Stored on the UICC/ISIM; one IMPI per UICC/ISIM

IMPU (*IP Multimedia Public User Identity*) (as the MSISDN at 5G level)

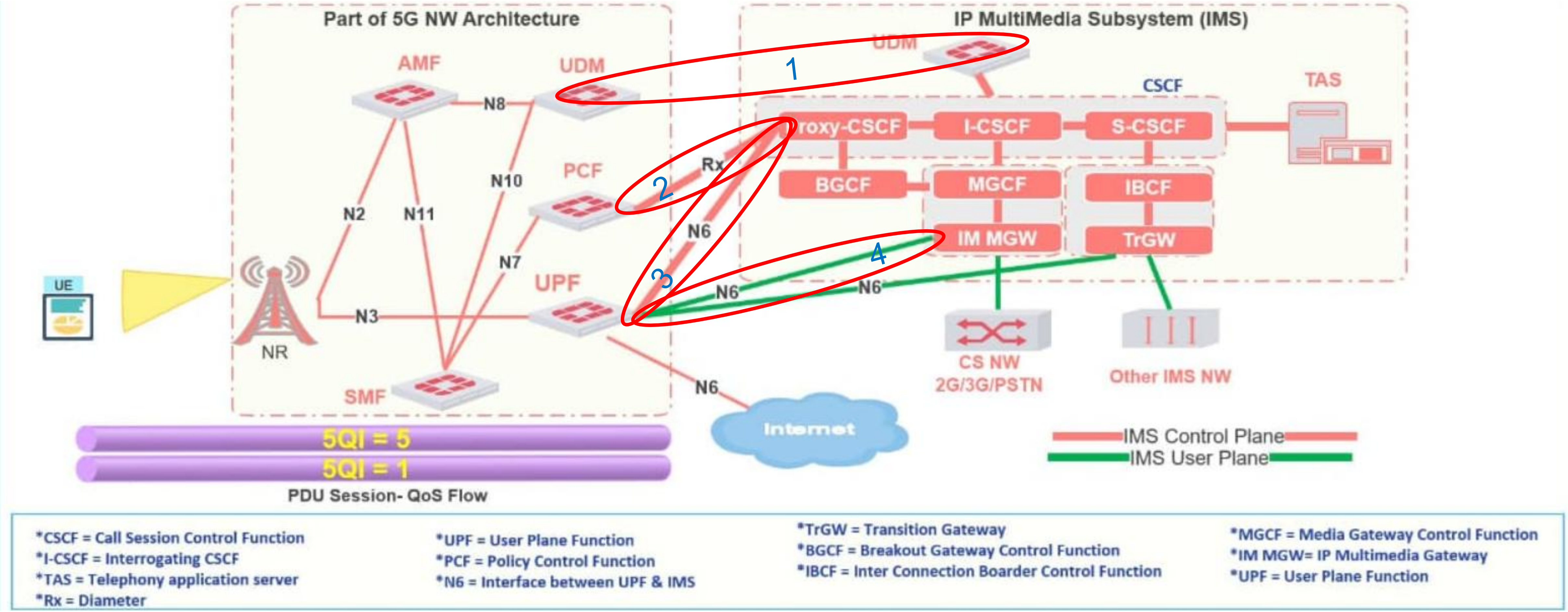
One or more **public identities used to identify a subscriber** over the public network

- **Usage**: Used for establishing communication sessions and is often presented as the caller ID
- **Format**: A *Universal Resource Identifier* (URI), such as a SIP URI or Tel URI (e.g., sip:john.smith@abc-company.se or tel:+31163279911)
- **Storage**: Stored on the UICC/ISIM; multiple IMPUs can be associated with one IMPI



IMS data is stored on the UICC within a dedicated application called the **ISIM (*IP Multimedia Services Identity Module*)**

5G to IMS interconnection



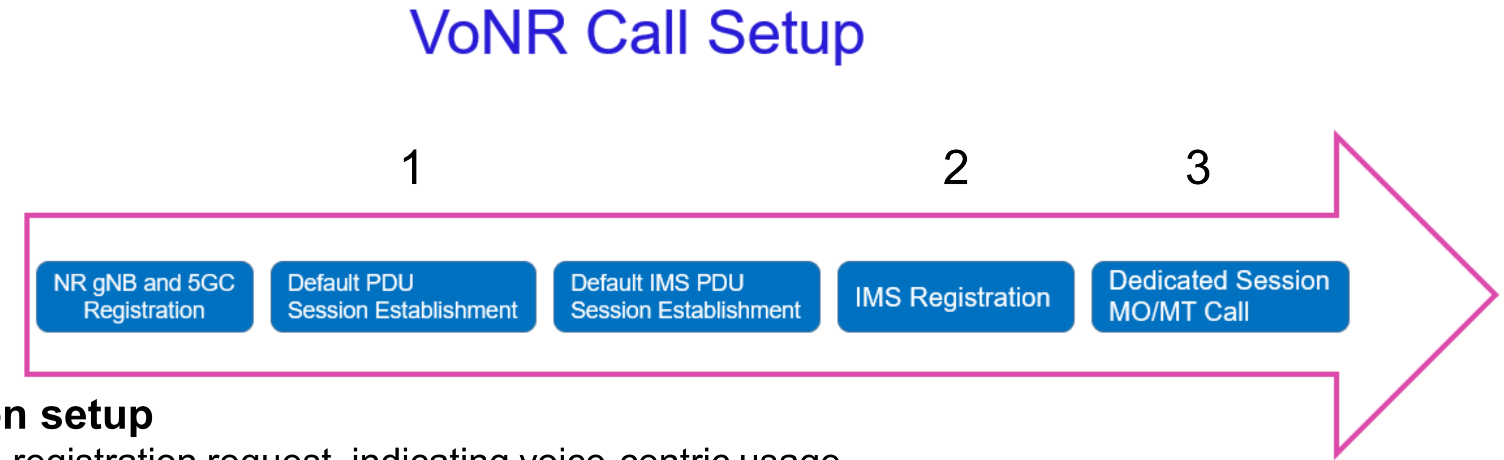
UDM as a common repository for 5G and IMS (1)

N5 (/Rx) interface between IMS P-CSCF (acting as AF) and 5G PCF (2)

Connectivity over a specific DNN ('IMS') in N6 interface

- IMS control plane (SIP) to P-CSCF (Gm) (3)
- IMS data plane to other terminating entity (4)

Voice over NR (VoNR)



1. Initial network registration and session setup

- The UE detects a 5G NR cell and performs a registration request, indicating voice-centric usage
- The UE establishes default PDU sessions:
    - One for general data (DNN = 'internet') with a non-GBR QoS flow (5QI = 6-9) and
    - Another for IMS signaling (DNN = 'ims') with a non-GBR QoS flow (5QI = 5; '***IMS Signalling***')

2. IMS registration and SIP negotiation

- The UE uses the dedicated IMS PDU session to register with the IMS using SIP messaging
- The UE and the IMS then perform SIP negotiation to agree on call parameters like codec, IP addresses, and port numbers

3. Voice call establishment

- Once SIP negotiation is complete, the 5GC sets up a dedicated PDU session with a GBR QoS flow (5QI = 1; '***Conversational Voice***') for the voice traffic for both the calling and called UEs
- The gNB sets up a dedicated DRB for this voice traffic

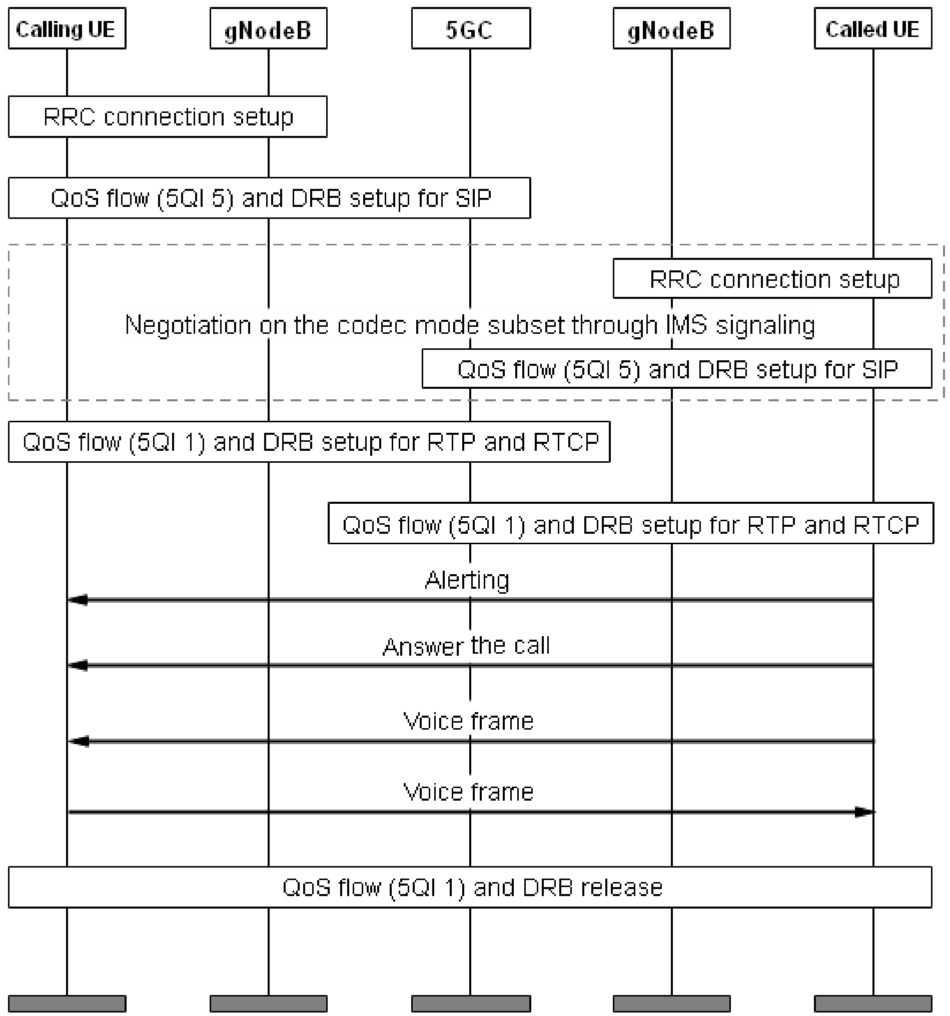
4. Voice data flow

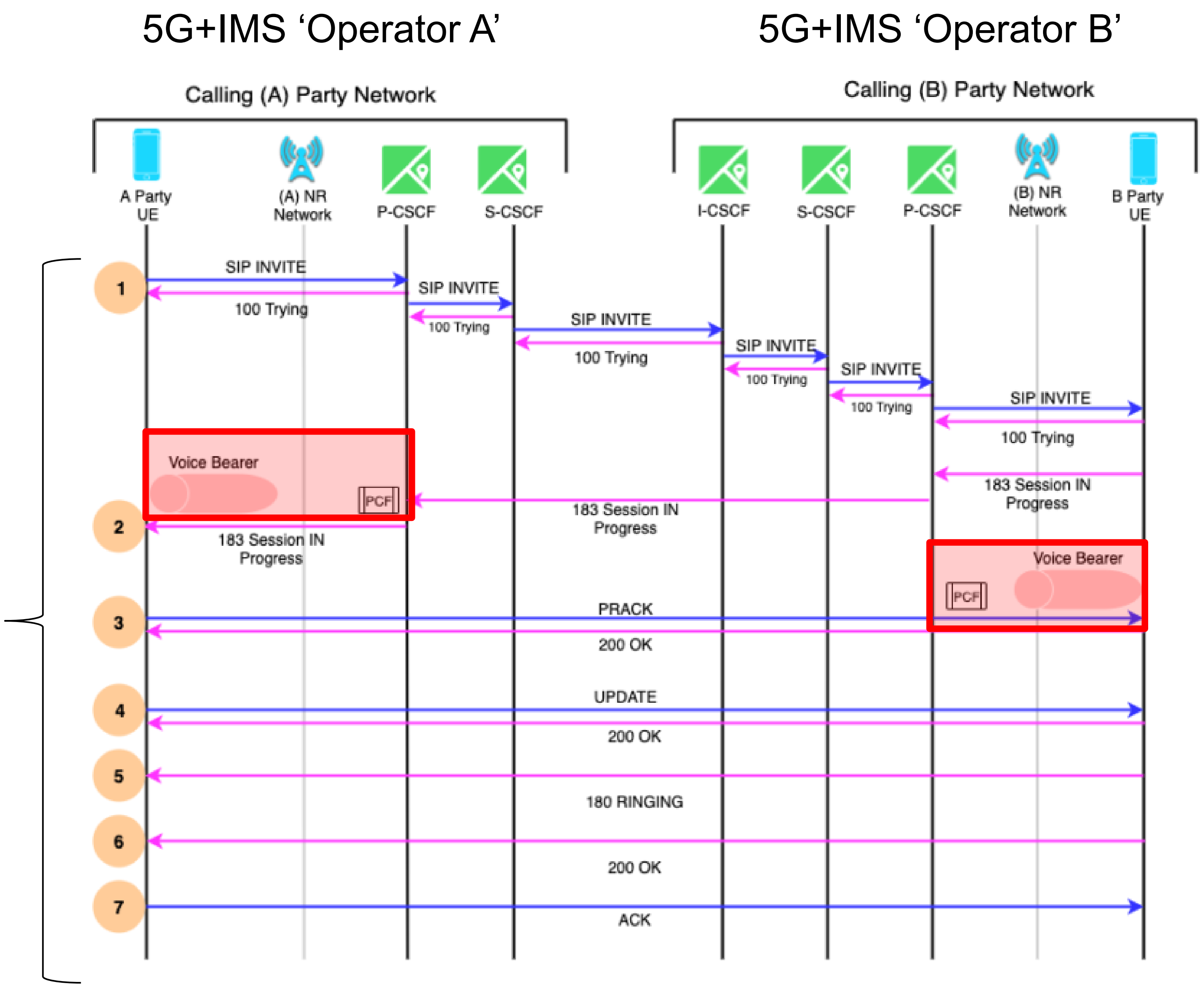
- Voice and control data (RTP and RTCP) flows through the dedicated PDU session and DRBs, using the established SIP parameters

5. Call termination

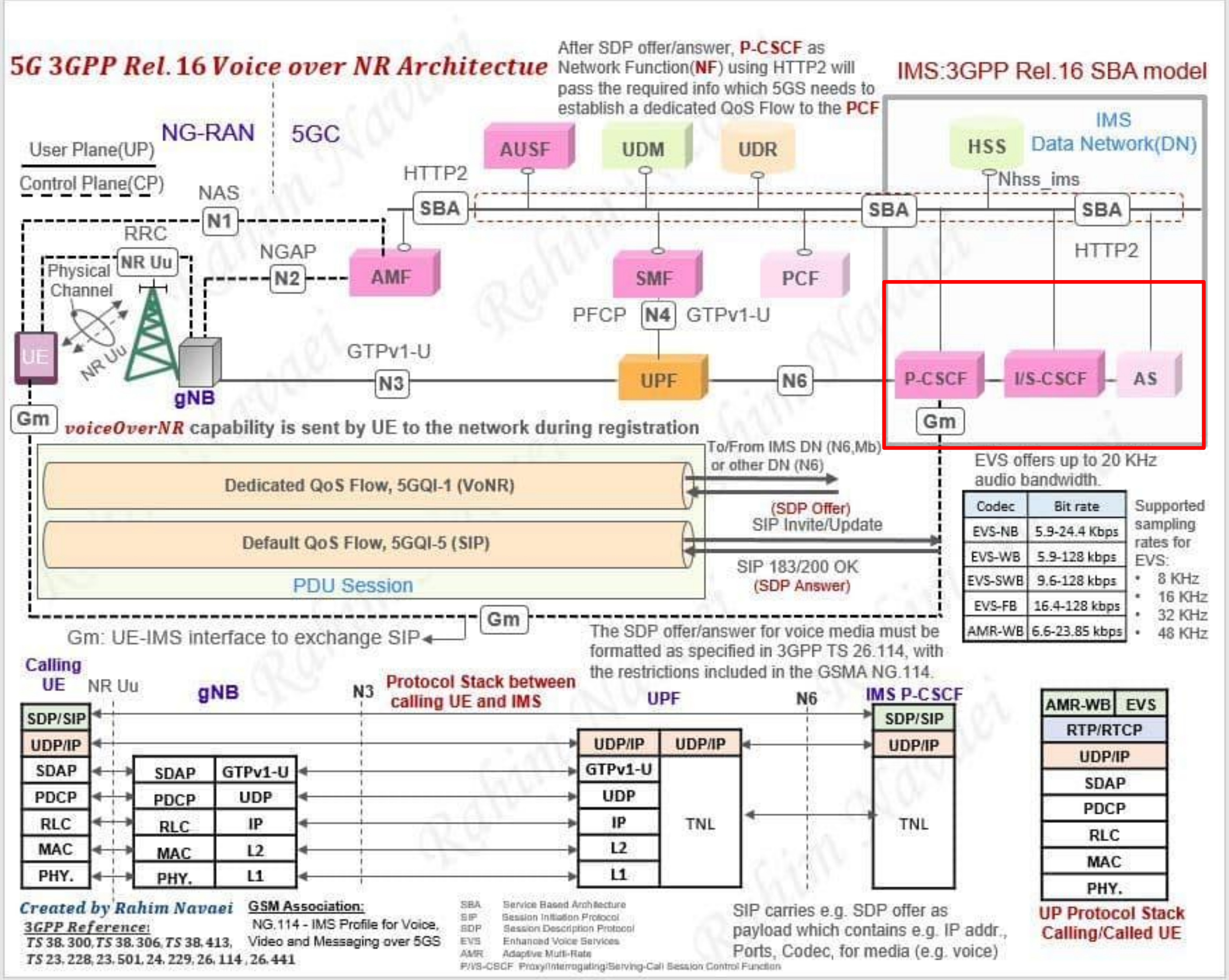
- When the call ends, the UEs release the dedicated QoS flows (5QI = 1), and the gNBs release the corresponding DRBs
- The default bearer for IMS (5QI = 5) is released later, typically when the UE enters an idle state

VoNR Call Flow





The 'alerting'/'ringing' notification is only send after the voice bearer (PDU Session with 5QI= 1) was possible (PCF triggered) in both directions of the conversation



5G 3GPP Rel.16 Voice over NR Architectue
After SDP offer/answer, P-CSCF as Network Function(NF) using HTTP2 will pass the required info which 5GS needs to establish a dedicated QoS Flow to the PCF
IMS:3GPP Rel.16 SBA model
NG-RAN
5GC
User Plane(UP)
Control Plane(CP)
NAS
N1
RRC
NR Uu
NGAP
N2
Physical Channel
NR Uu
UE
gNB
HTTP2
SBA
AUSF
UDM
UDR
AMF
SMF
PCF
PFCP
N4
GTPv1-U
GTPv1-U
N3
UPF
N6
IMS
Data Network(DN)
HSS
Nhss_ims
SBA
SBA
HTTP2
P-CSCF
I/S-CSCF
AS
Gm
Gm
voiceOverNR capability is sent by UE to the network during registration
To/From IMS DN (N6,Mb) or other DN (N6)
Dedicated QoS Flow, 5GQI-1 (VoNR)
Default QoS Flow, 5GQI-5 (SIP)
PDU Session
(SDP Offer)
SIP Invite/Update
SIP 183/200 OK
(SDP Answer)
Gm
Gm: UE-IMS interface to exchange SIP
EVS offers up to 20 KHz audio bandwidth.
Codec | Bit rate
EVS-NB | 5.9-24.4 Kbps
EVS-WB | 5.9-128 kbps
EVS-SWB | 9.6-128 kbps
EVS-FB | 16.4-128 kbps
AMR-WB | 6.6-23.85 kbps
Supported sampling rates for EVS:
• 8 KHz
• 16 KHz
• 32 KHz
• 48 KHz
The SDP offer/answer for voice media must be formatted as specified in 3GPP TS 26.114, with the restrictions included in the GSMA NG.114.
Calling UE
NR Uu
gNB
N3
Protocol Stack between calling UE and IMS
UPF
N6
IMS P-CSCF
SDP/SIP
UDP/IP
SDAP
PDCP
RLC
MAC
PHY.
SDAP | GTPv1-U
PDCP | UDP
RLC | IP
MAC | L2
PHY. | L1
UDP/IP | UDP/IP
GTPv1-U
UDP
IP
L2
L1
TNL
SDP/SIP
UDP/IP
TNL
AMR-WB | EVS
RTP/RTCP
UDP/IP
SDAP
PDCP
RLC
MAC
PHY.
UP Protocol Stack Calling/Called UE
Created by Rahim Navaei
3GPP Reference:
TS 38.300, TS 38.306, TS 38.413, TS 23.228, 23.501, 24.229, 26.114, 26.441
GSM Association:
NG.114 - IMS Profile for Voice, Video and Messaging over 5GS
SBA Service Based Architecture
SIP Session Initiation Protocol
SDP Session Description Protocol
EVS Enhanced Voice Services
AMR Adaptive Multi-Rate
P/I/S-CSCF Proxy/Interrogating/Serving-Call Session Control Function
SIP carries e.g. SDP offer as payload which contains e.g. IP addr., Ports, Codec, for media (e.g. voice)

AI/ML

Areas of AI/ML in Mobile Networks

Context:

- Mobile networks are increasingly more complex
- Low ROI (*Return on Investment*) and Decreasing ARPUs (*Average Revenew Per User*) puts pressure in investments
- Availability and efficiency and must be very high

→Common OSS/BSS tools no longer cope with emerging challenges

Adopt and integrate AI&ML in the network and services for:

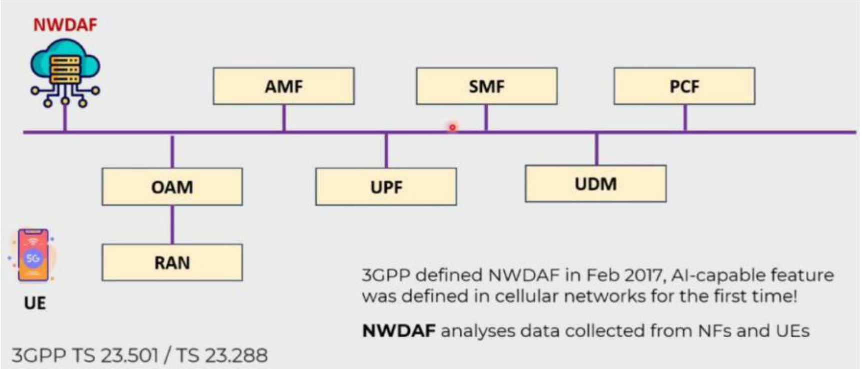
- Increased efficiency
- Cost reduction
- New revenue streams
- Net operation optimization
- Predictive maintenance
- Customer care
- …

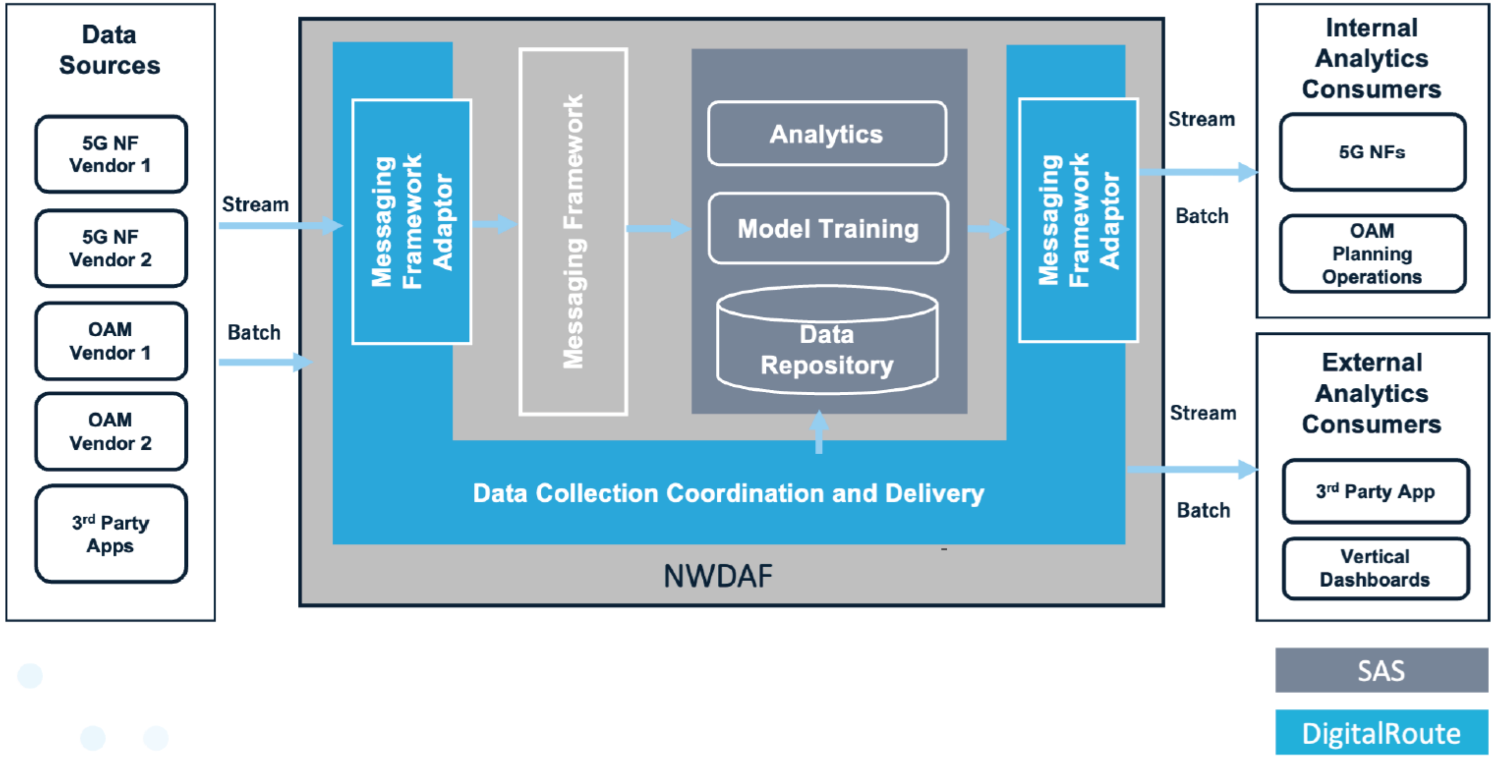
3GPP Standardization timeline

Rel-15	Rel-16	Rel-17	Rel-18	Rel-19 (on going)
Architecture • NWDAF **Use Cases** • Slice load level START	**Architecture** • Intelligent Network architecture **Use Cases** • Service Experience • NF load • Network performance • UE mobility, communication and abnormal monitoring • Congestion • QoS sustainability	**Architecture** • Decoupling NWDAF into AnLF and MTLF • Multi NWDAFs Layered Deployment • Data collection and storage efficiency improvement **Use Cases** • Dispersion • WLAN and DN performance • Session management congestion • Redundant transmission	**Architecture** • Correctness of analytics and ML models enhancements • 5GC Horizontal Federal Learning • Analytics exposure in roaming **Use Cases** • PFD Determination • Location accuracy • Relative proximity • Movement Behaviour	**Architecture** • Vertical Federated Learning **Use Cases** • LCS to support Direct AI/ML based Positioning • NWDAF-assisted policy control and QoS enhancement • NWDAF supports network abnormal behaviours mitigation and prevention

Network Data Analytics Function (NWDAF)

Responsible for collecting, analyzing, and generating insights from network data to enable intelligent automation and optimization





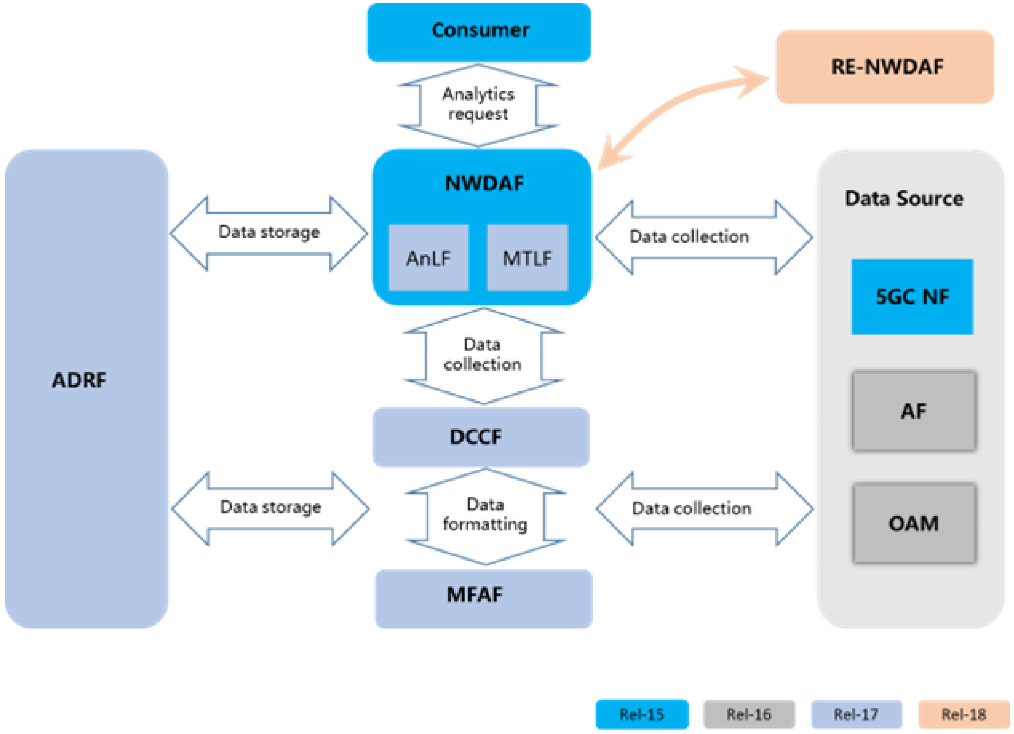
3GPP Use-cases

1. Slice load level analytics
2. Observed Service Experience
3. NF load analytics
4. Network Performance analytics
5. UE Mobility analytics
6. UE Communication analytics
7. User Data Congestion analytics
8. QoS Sustainability analytics
9. Dispersion analytics
10. DN Performance analytics
11. WLAN Performance analytics
12. Session Management congestion control
13. Redundant Transmission Experience
14. Abnormal behavior
15. User data congestion

NWDAF

3GPP TS 28.104

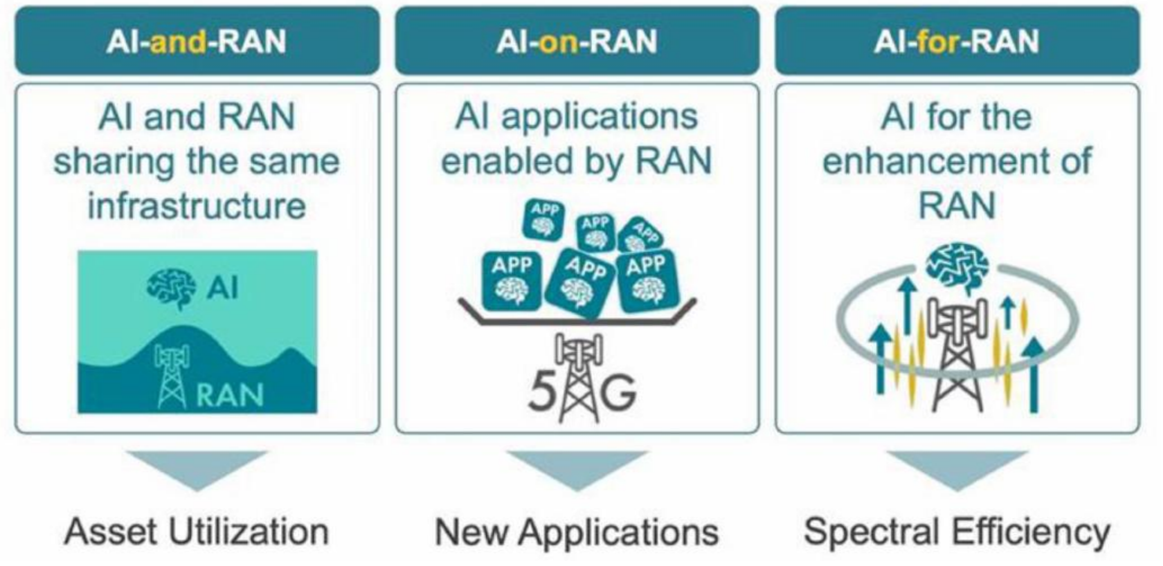
- Data collection based on **subscription to events** provided by AMF, SMF, UPF, PCF, UDM, NSACF, AF (directly or via NEF) and OAM
- Analytics and Data collection using the DCCF (*Data Collection Coordination Function*)
- Retrieval of information from data repositories (e.g. from UDR via UDM for subscriber-related information or optionally via NEF(PFDF) for PFD information)
- Data collection of location information using LCS (finer granularity location information determined by LMF)
- Storage and retrieval of information from ADRF (*Analytics Data Repository Function*)
- Analytics and Data collection from MFAF (*Messaging Framework Adaptor Function*)
- Retrieval of information about NFs (e.g. from NRF for NF-related information)
- On demand provision of analytics to consumers
- Provision of bulked data related to Analytics ID(s)
- Provision of Accuracy information about Analytics ID(s)
- Provision or retrieval of ML Model Accuracy Information or ML Model accuracy degradation about a ML Model
- Federated and Vertical Federated Learning

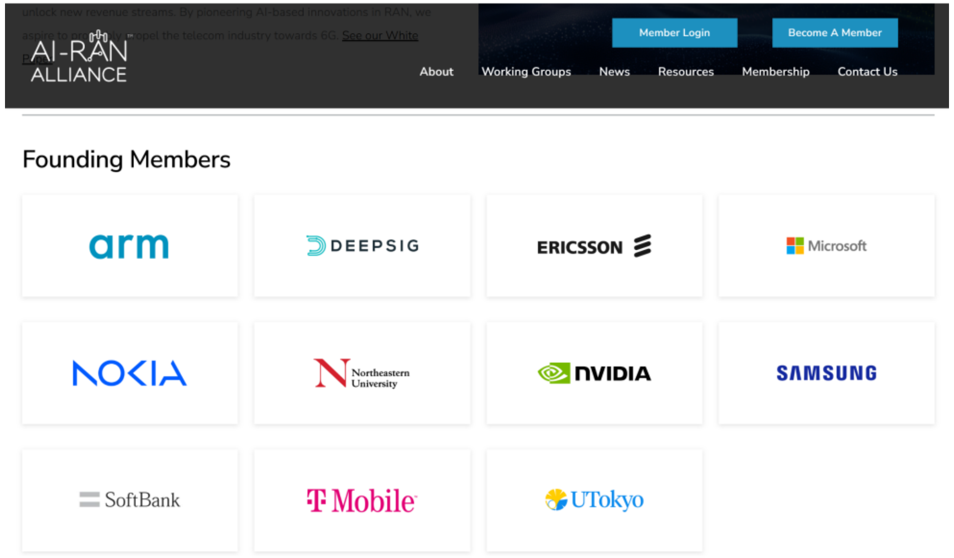


ADRF: *Analytics Data Repository Function*
AnLF: *Analytics Logical Function*
DCCF: *Data Collection Coordination Function*
MFAF: *Messaging Framework Adaptor Function*
MTLF: *Model Training Logical Function*
OAM: *Operations, Administration, and Maintenance*

AI at 5G RAN level

AI RAN Alliance: https://ai-ran.org/



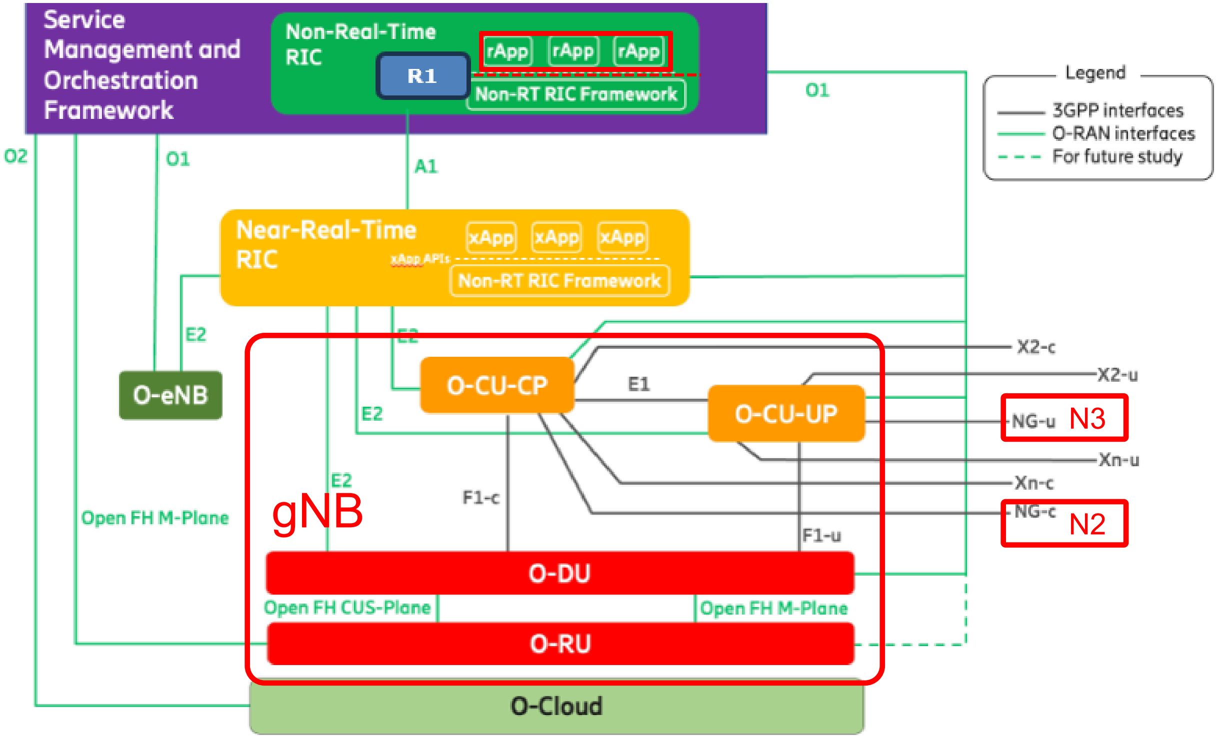


Working Groups	Focus	Key Benefits
AI and RAN	Running both AI and RAN workloads on a common, shared computing infrastructure, such as a GPU cloud platform	Maximized resource utilization, reduced total cost of ownership (TCO), and creation of new AI-driven revenue streams by utilizing otherwise idle hardware
AI on RAN	Deploying AI services at the network edge, leveraging RAN capabilities to enhance AI application performance and capabilities	Increased operational efficiency, reduced backhaul traffic, enablement of new consumer and enterprise edge services, and real-time localized inferencing
AI for RAN	Embedding AI/ML models into the radio signal processing layer to enhance RAN capabilities directly	Improved spectral efficiency, better radio coverage and capacity, optimized latency, and predictive maintenance

ORAN and AI

Open RAN (ORAN) mission: re-shape RANto be more intelligent, open, virtualized and fully interoperable

- Inclusion of the *Near-Real-Time* and the *Non-Real-Time RICs*: Automate and optimize the RAN, reducing the mobile operator's TCO (Total Cost of Owership)
- AI and ML applications running as xApps and rApps in the Near- and Non-Real Time RICs are a key architecture component of O-RAN



rApp: RAN Intelligent Controller applications

Run on the Non-Real Time RIC (Non-RT RIC)

Handling network management and optimization tasks that don't require immediate, real-time responses:

- Traffic Steering
- Energy Savings
- Anomaly Detection
- Long-term Load Balancing
- Mobility Optimization
- Network Planning and Deployment
- Intelligent Automation

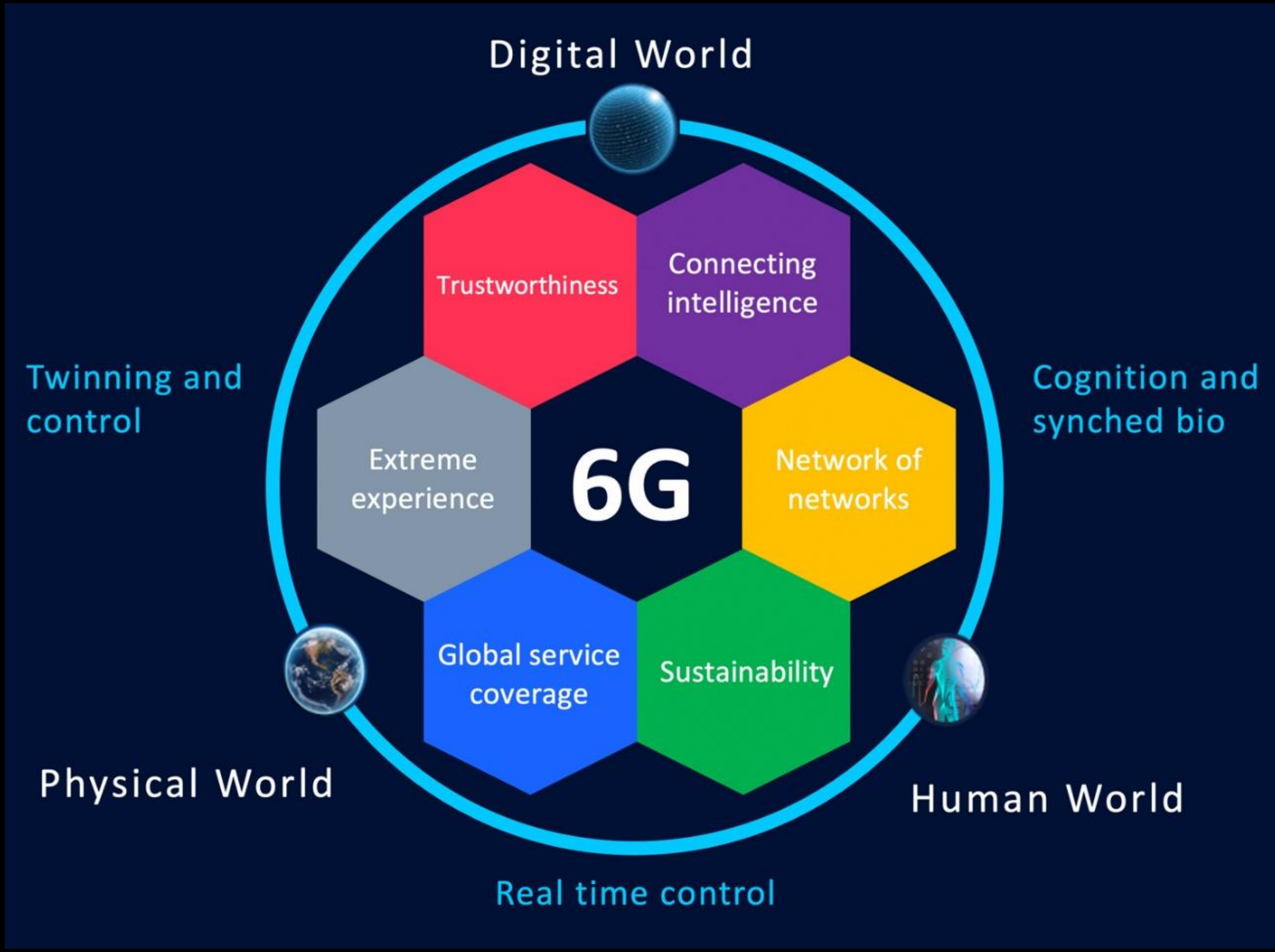
Communicate with the non-RT RIC via the **A1 interface**:

- specific application protocol that has been enhanced to support A1 policies and for the Non-RT-RIC to provide enrichment information to the Near-RT-RIC

“While 5G has enabled us to consume digital media anywhere, anytime, the technology of the future should enable us to embed ourselves in entirely virtual or digital worlds.

In the world of 2030, human intelligence will be augmented by being tightly coupled and seamlessly intertwined with the network and digital technologies.”

Hexa-X Consortium



• Beyond 5G, towards 6G

"The 6G network is likely to be a concept, a virtual one, and not a "real" network you can put a boundary around", Roberto Saracco, EIT Digital

Perceived unlimited bandwidth with unperceived latency

Technological pillars

- Operate at higher radio frequencies (Thz)
- Inclusive of all sort of access technologies, expanding to:
  - Non Terrestrial Networks (NTN), Optical Wireless Communications (OWC) and Large Intelligent Surfaces (LIS)
- Decentralized, flatter network with a stronger edge role, for close, distributed cloud services
- Increased direct devices interactions
- Artificial Intelligence presence at all layers, with cross-domain interactions
- Higher security, secrecy and privacy

Societal and economic impacts

- Ubiquitous connectivity, powered by wireless communications (radio and optics)
- Richer set of connected devices

Enabling:

- The smartphone disaggregation
- Multisensorial interactions
  - Via Brain-Computer Interactions (BCI), smart body implants and eXtended Reality (XR) devices
- Massification of Machine Type Communications (MTC)
  - Connected Robotics and autonomous systems