Research on Cybersecurity Threats and Solutions in RATs and C-RAN 5G Network

MINT-709
Capstone Project Report

Presented by:
Harsh Vardhan

University of Alberta
Master of Science in Internetworking
Department of Electrical and Computer Engineering

Supervisor
Sandeep Kaur

ABSTRACT

Alexander Graham Bell made the first successful practical telephone conversation on March 10, 1876, marking a significant milestone in telecommunication history. Since then, the concept of telephone communication has expanded, and Motorola introduced the first mobile phone in 1973. In Japan, 1G or first-generation analog networks appeared in 1979, 2G or second-generation digital cellular networks based on GSM technology occurred in 1991, 3G or third-generation wireless mobile telecommunications technology based on UMTS and CDMA appeared in 2001, and 4G or fourth-generation based on LTE standard appeared in 2009.

With the growing number of mobile connections and the continual rise of mobile data, mobile usage will be different than it is now. 5G is a next-generation network that provides faster data speeds, ultra-low latency, and an exceptionally reliable system. However, it is challenging for service providers to maintain uninterrupted service in densely populated areas. The carriers install more base stations in small areas to maintain a smooth traffic flow. Instead of meeting the enormous demand, these base stations generate interference with one another, resulting in energy loss and increased costs. Even during non-peak hours, they operate at maximum capacity, wasting energy.

As a result, Centralised RAN must be implemented. The base station consists of two components in CRAN, the Remote Radio Head and the Base Band Unit. The BBUs are grouped in a centralized pool and connected to the RRH via front haul. The BBUs in the BBU pool share their resources and energy. It can reduce the cost of network deployment and operation while also improving energy efficiency, mobility, and coverage.

On the other hand, CRAN technology in 5G mobile communication networks needs to deal with several security challenges relating to virtual network functions and software-defined networking. This project aims to study and present a comprehensive evaluation of existing security studies in the field of C-RAN and corresponding security threats and attacks. Furthermore, we indicate open solutions research issues and propose future research trends.

ACKNOWLEDGEMENT

I am pleased to express my gratitude to my mentor, Ms. Sandeep Kaur, who guided me throughout the project and provided invaluable motivation and suggestions to expand my knowledge base. In addition, she offered insight and freedom to work on my project while ensuring that I stayed on track and did not stray from my project's core.

I would also like to express my sincere gratitude to my professors, **Dr. Mike McGregor** and **Mr. Shahnawaz Mir,** for providing me with assistance and such a great opportunity.

Last but not least, I would like to express my gratitude to my classmates, professors, and the University of Alberta for assisting and supporting me in achieving this goal whenever feasible.

Table of Contents

Table of Figures

List of Tables

1. Introduction

1.1 Evolution Of Wireless Technologies

Since Guglielmo Marconi first showed radio's ability to establish continuous communication with ships, the ability to communicate with people on the move has advanced dramatically. Mobile wireless evolved in a short period.

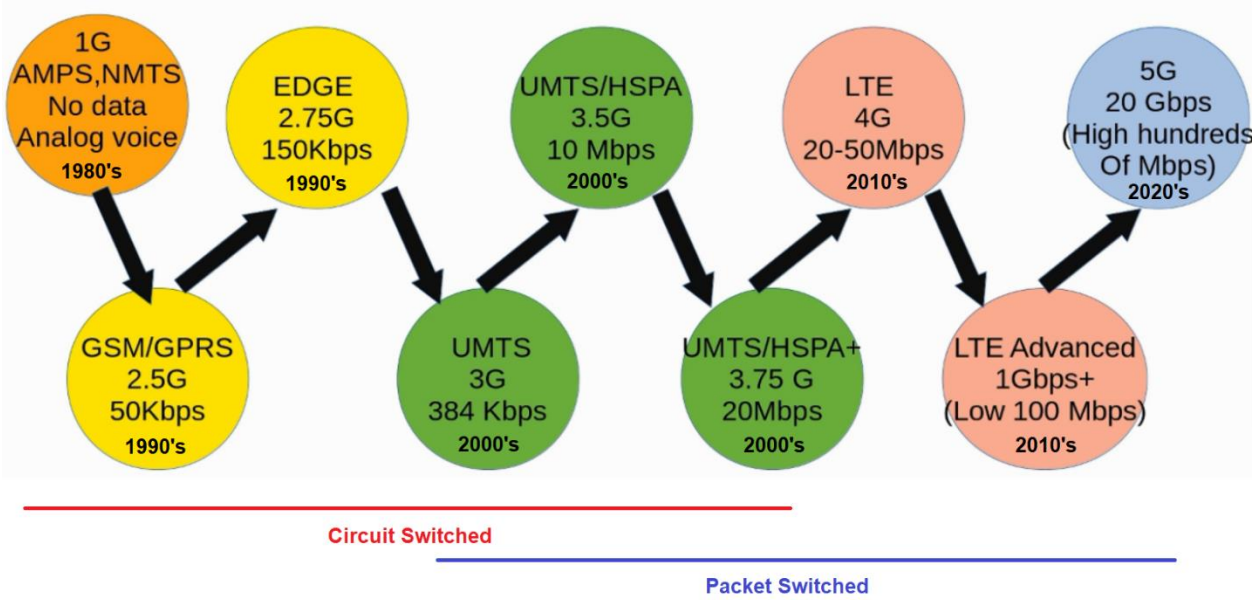


Fig. 1. Evolution of wireless network.

As shown in Fig.1, the mobile network has evolved considerably in the past few decades. The majority refers to improvements in definition, speed, infrastructure, bandwidth, data size, latency, and other aspects of mobile wireless generation. The first-generation communication was through the handheld analog system for phone calls without a wired network. Second-generation technology includes digital systems and text messaging. Third-generation cell technology allowed faster data processing, improved performance, and increased digital media capacity. The fourth-generation combines 3G and broadband services to enable mobile networks to function and overcome previous restraints and limitations. For example, there was more bandwidth at a lower cost of power. The current fifth generation provides users with the most up-to-date technology, allowing them to communicate at higher data rates and more dependably and securely than the previous fourth generation.

1.2 1G Mobile Communication System

First-generation mobile systems used analog transmission for voice services and were developed based on analog technologies. This technology used analog signals to communicate voice calls with a speed of up to 2.4 kbps. In 1979, Nippon Telephone and Telegraph (NTT) in Tokyo, Japan, developed the first cellular operational system in the world. Europe also came forward with popular analog systems were Nordic Mobile Telephone (NMT) and Total Access Communication System (TACS); other analog systems were also introduced in the 1980s across Europe. In the United States, the Advanced Mobile Phone System (AMPS) was launched in 1982 and was most widely used. The Advanced Mobile Phone system was allocated a 40 MHz bandwidth within the 800-900 MHz frequency range by the Federal Communication Commission (FCC). The AMPS used Frequency Division Multiplexing (FDMA) as the modulation technique and could scale up to a 7-cell cluster for frequency reuse. The system used two channels, one for forward and one for backward communication.

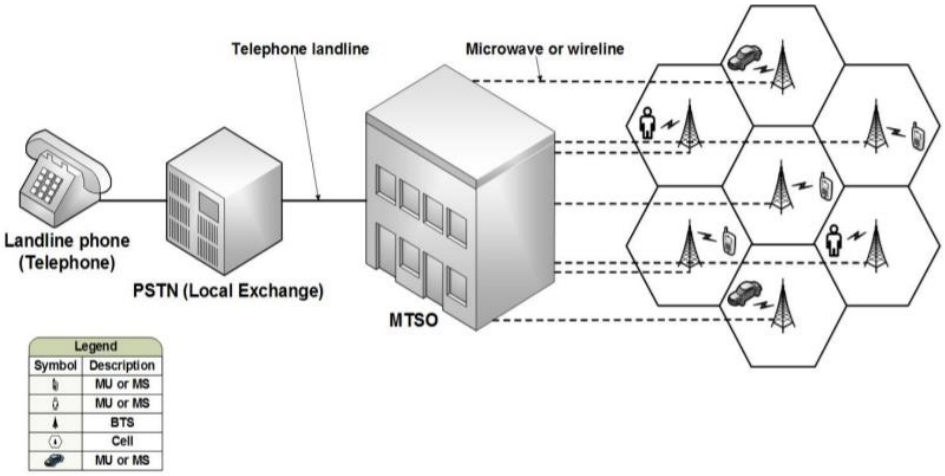


Fig. 2. 1G AMPS Architecture [1]

1.2.1 Features

1. Used only Analog systems
2. Offered only voice calls
3. Low capacity due to FDMA multiplexing
4. Made mobile communication possible

1.2.2 Disadvantages

1. Weak security (easy eavesdropping)
2. No roaming between different operators
3. The voice quality was poor
4. Poor handoffs
5. Limited capacity and required large phone size

1.3 Second-Generation (2G)

Second-generation (2G) mobile systems were introduced in 1991 in Finland based on the GSM standard. It brought digitalization to communication systems and was the first to offer data service, better security, and efficient use of spectrum available. 2G provided text, message, and multimedia services (MMS). Moreover, all text messages were digitally encrypted, intended only for the receiver. The first digital system was IS-54 (TDMA), IS-95 (CDMA) in 1993, and IS-136 in 1996.

2G systems used multiple access schemes such combination of Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA), and Code Division Multiple Access (CDMA) for more efficient bandwidth/spectrum allocation.

1.3.1 GSM

The European Telecommunications Standards Institute (ETSI) initiated the development of the first version of Global System for Mobile. GSM operates in the following frequency band ranges, i.e., GSM-900, GSM-1800, and GSM-1900 with 124, 374, and 299 radio channels, respectively. Time-division multiple access (TDMA) is used to multiplex up to 8 calls per channel in the 900 and 1800 MHZ bands in GSM. The frequency ranged from 30 to 200 kHz and allocated one channel for uplink and one for downlink transmission. As a result, GSM delivers voice, and circuit-switched data speeds up to 14.4kbps.

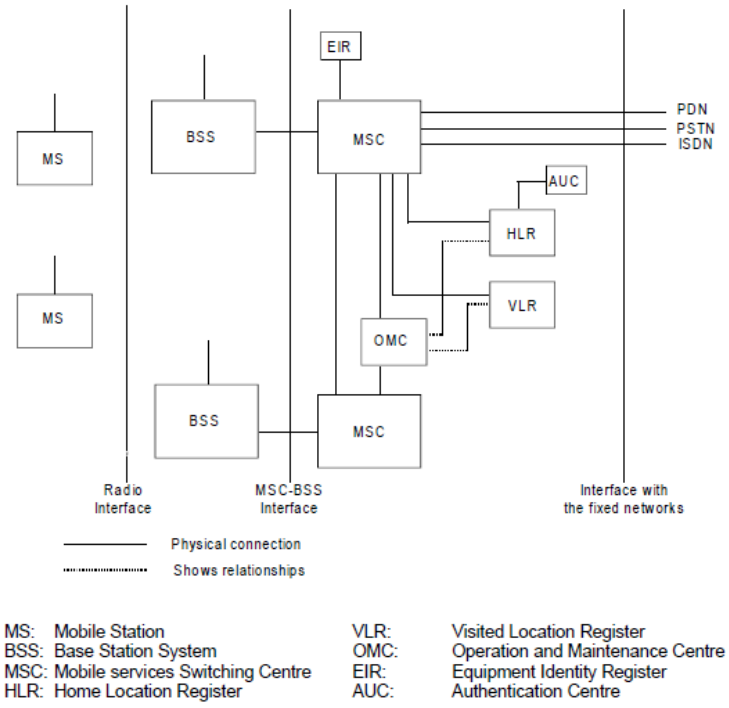


Fig. 3. 2G GSM architecture [2]

1.3.2 2.5 G

GPRS is, also known as 2.5G, is an extension of the existing 2G GSM network and proposed to provide the capacity of launching packet-based services and increased data rates. The central ideology was to deliver IP services without replacing the network infrastructure as the demand for both voice and data increased. To improve data rates, it used different coding schemes CS1, CS2, CS3, and CS4. Data transmission rate had a minimum speed of 50 kbps and 384 Kbps. In addition, serving GPRS Support Node (SGSN) and gateway GPRS Support Node (GGSN) is introduced at the core network.

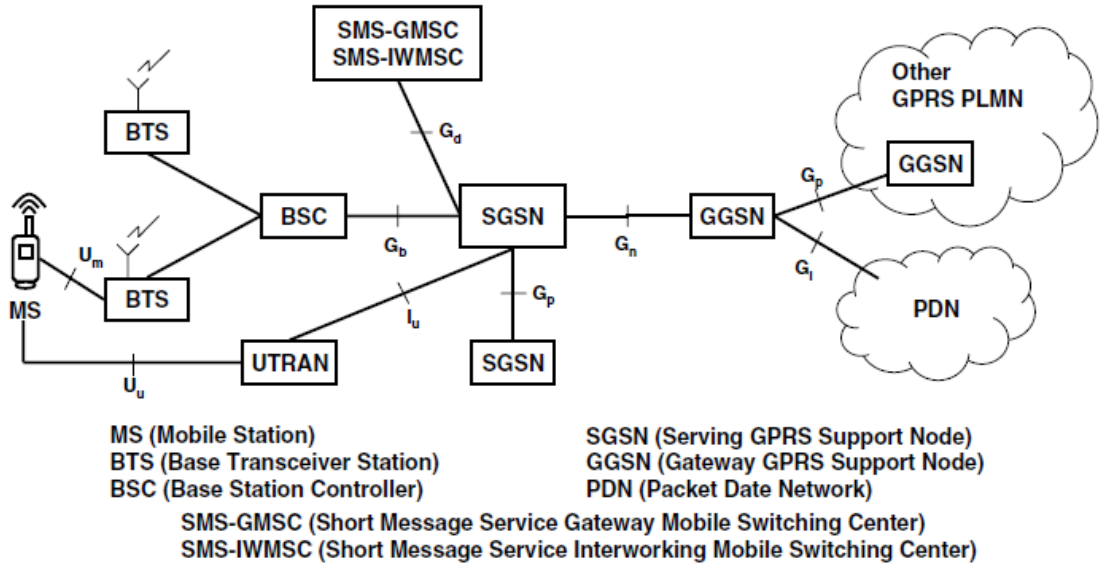


Fig. 4. GPRS system architecture [3]

1.3.3 2.75 G

In 2003 significant improvements were made to the existing GSM systems with the deployment of Enhanced Data rates for GSM Evolution (EDGE) networks. A backward-compatible digital mobile phone technology is also known as IMT Single Carrier (IMT-SC). The system used nine different modulation and coding schemes (MCS- 8PSK) to enhance the data rates. As a result, the data rates are increased to 384 kbps for downlink and 60 kbps for uplink, respectively.

1.3.4 Features

1. Digital data services introduced (SMS, email)
2. Higher transmission rates
3. Backward compatible mobile technology

1.3.5 Disadvantages

1. Weak signals in less populated areas
2. Unable to handle complex data (Video)

1.4 Third-Generation (3G)

The third generation of mobile communication is 3G, and technology succeeds 2G due to the rapid increase in data services and Internet Protocol demand. This technology was designed to provide performance at high speed. 3G is not one standard; it is a culmination of criteria that can all work together. Such demand and advancement led ETSI to initiate a strategy and develop a new Universal Mobile Telecommunications System (UMTS) system. Meanwhile, the International Telecommunication Union (ITU) recommended the 3G system known as International Mobile Telecommunications 2000 (IMT-2000) and assigned a communication spectrum between 400 MHz and 3 GHz. The critical mechanism used for medium access in 3G is CDMA which allows multiple mobile stations to transmit simultaneously and in the same frequency band.

Wireless Generation	Systems	General Service	Comments
Third (3G)	CDMA2000/ WCDMA	Packet Data and Voice services. High-speed data and voice.	Defined by IMT 2000. Europe (UMTS – WCDMA) America (UMTS / CDMA2000) Asia (UMTS / CDMA2000)

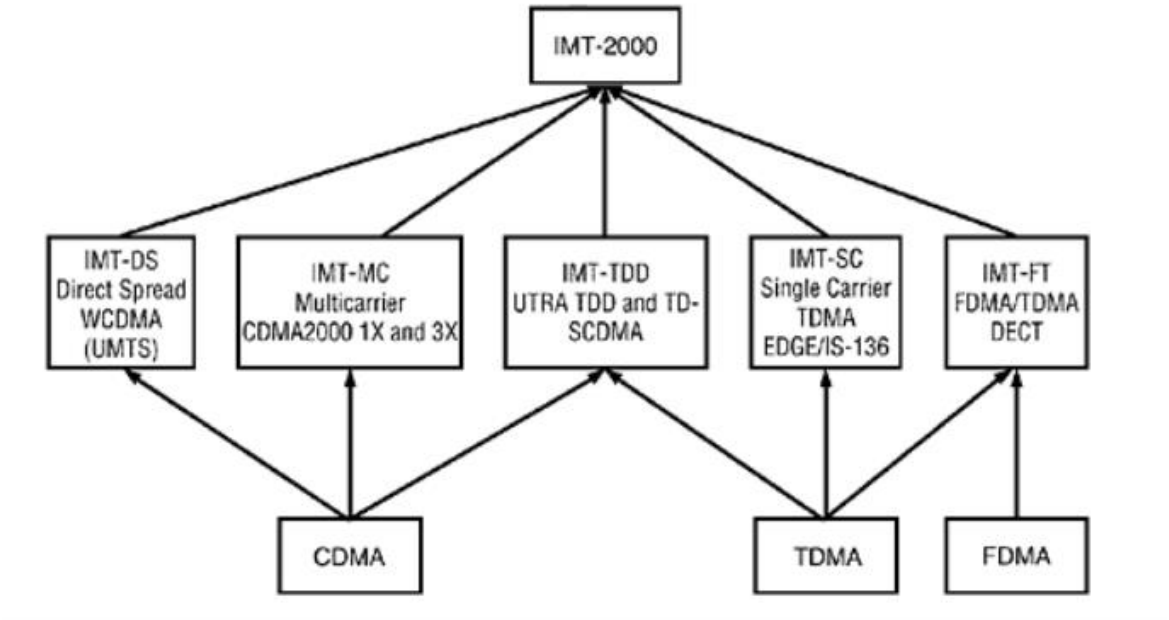


Fig. 5. IMT2000 [4]

IMT2000/3G can be summarised as follows [4] :

- A range of technologies provides different frequency bands, channel bandwidths, and modulation formats.
- No single platform, application, or technology for 3G infrastructure exists.
- 3G high-speed data is often used in wireless mobile and stationary applications. For example, IMT-2000 proposes data speeds of 144 Kbps at moving rates, 384 Kbps for outdoor, and 2 Mbps for indoor.

The establishment of the Third Generation Partnership Project (3GPP) became the standardization of cellular networks with its initial release as 3GPP release 1999 consisting of GSM specification enhancement.

The main components include

- BS (Base Station) or NodeB
- RNC (Radio Network Controller)
- WMSC (Wideband CDMA Mobile Switching Centre)
- SGSN/GGSN.

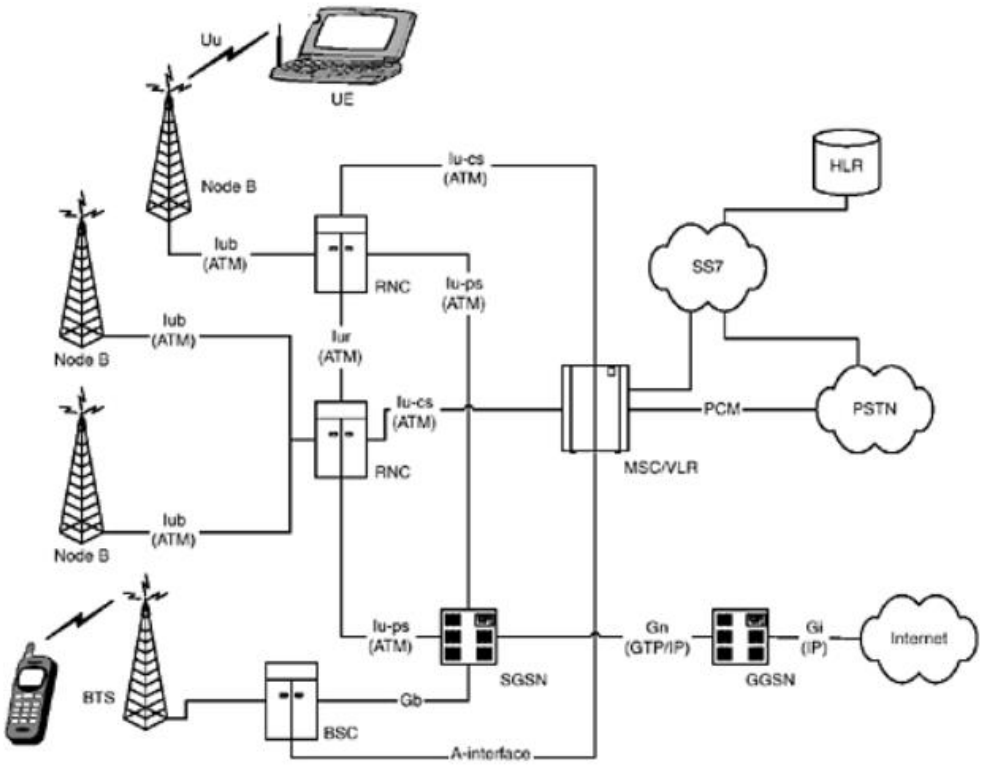


Fig. 6. 3GPP Release 1999 Network Architecture [4]

1.4.1 Features

- Increased bandwidth and data transmission rates
- Immense power & broadband capacity
- Video-conferencing support
- Internet TV

1.4.2 Drawbacks

- Require special hardware and up-gradation, which is costly
- Complex modulation/demodulation increased high-power consumption.
- Roaming

1.4.3 3.5 G – High-Speed Downlink Packet Access

It consists of two mobile protocols by 3GPP, High-Speed Downlink Packet Access (HSDPA) in Release 6 and High-Speed Uplink Packet Access (HSUPA) in Release 7 that improves the existing 3G performance using WCDMA.

- Increased peak data rates : 14 Mbit/s (downlink) and 5.76 Mbit/s (uplink)

1.4.3.1 Evolved High-Speed Packet Access +

- Increased data rates up to 337.5 Mbit/s by adding 64QAM modulation, MIMO operations

1.5 Fourth Generation (4G)

The growing demand for a higher data rate for Internet access via mobile phones brought forward the fourth generation of the cellular network to deliver broadband data transmission and broadcasting for a high volume of users. NTT DoCoMo conducted the first successful 4G field trials to achieve 1 Gbps downlink in Tokyo, Japan, in 2005.

ITU's Radio communications Sector (ITU-R) in 2008 started developing a new system known as IMT-Advanced. The main requirements for this new system include supporting 100 Mbps and 1 Gbps peak data rate for high and low mobility scenarios, respectively, bandwidth scalability up to 100 MHz, mobility support with up to 350 km/h, and improved spectral efficiency [5]. As a result, the 3GPP develops the technology and standard in Release 8.

1.5.1 LTE

LTE is sometimes known as 3.95G; as it does not meet the requirements for 4G as specified by ITU, it is often marketed as a 4G technology. LTE uses Multi carrier CDMA or OFDM (Orthogonal Frequency Division Multiplexing) mechanisms for medium access. In addition, LTE supports FDD and TDD operations and provides flexible symmetric and asymmetric spectrum functions.

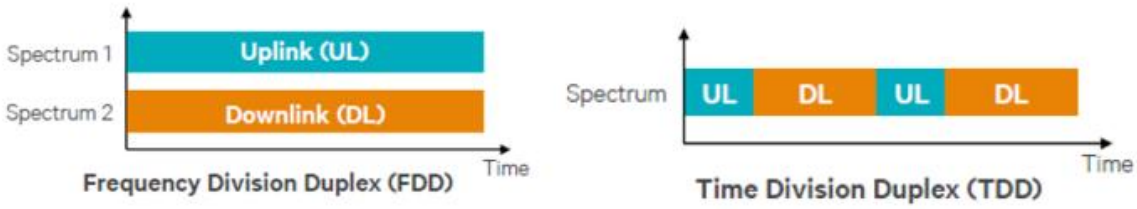


Fig. 7. Uplink and Downlink Spectrum in FDD and TDD

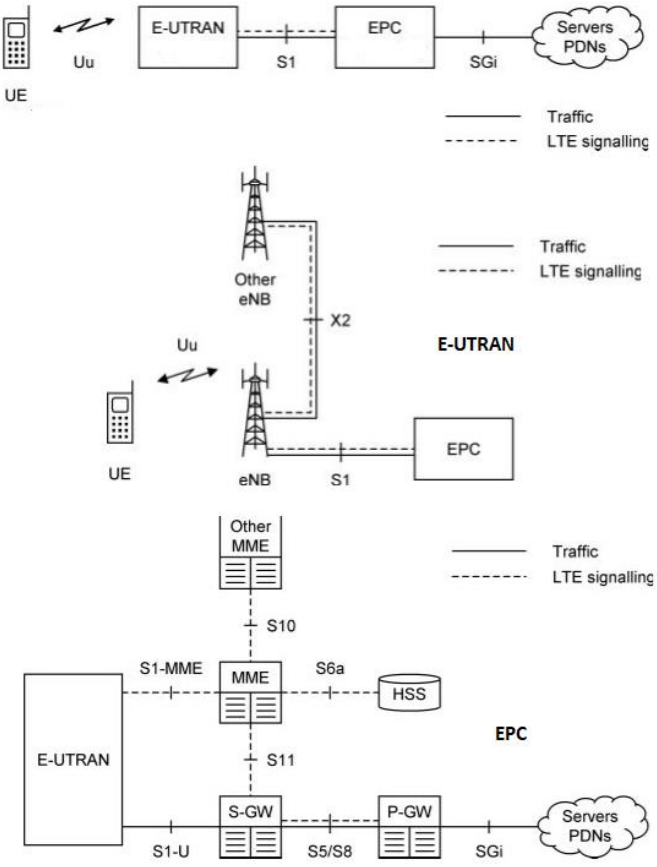


Fig. 8. LTE architecture [5]

From an architectural standpoint, the main components of a 4G LTE system are :

1.5.2 Evolved Node B (eNB)

- Radio transmission and reception
- Radio resource management
- Admission control and Resource block allocation
- Encryption of user data and signaling
- Coordination for handover with other eNB

1.5.3 Mobile Management Entity (MME)

- Manages UE registration and authentication
- Stores UE temporary data
- Bearer management and establishment

1.5.4 Serving Gateway (S-GW)

- Handles user data connectivity
- Packet routing and forwarding
- Inserting QoS markers into IP headers
- Data path anchor
- Lawful interception

1.5.5 Packet Data Network Gateway (P-GW)

- IP address allocation during bearer establishment
- Filter malware and unauthorized traffic
- Inserting QoS markers into IP headers

1.5.6 Policy And Charing Rules Function (PCRF)

- EPS bearer creation
- Formulates QoS policy and charging rules for data flow

1.5.7 Home Subscriber Server (HSS)

- The central database for subscriber information
- Access authorization
- Tracking UE

Multicasting and interference mitigation are significant enhancements in the 4G era. Energy efficiency is also one of the critical considerations by 3GPP Release 12/13 for green computing. LTE for machine-type communication (MTC), M2M, and IoT.

1.5.7.1 Features :

- Maximum 100 Mbps downlink and around 50 Mbps uplink speed
- Ease of access and less latency
- Improved voice and HD streaming quality
- Interoperability and Easy Roaming

1.5.7.2 Drawbacks :

- New CAPEX/OPEX cost involved.
- No backward compatibility with 3G
- Higher data service cost.

1.5.8 LTE Advance

The Third Generation Partnership Project (3GPP) had presented the foundations of the future Long Term Evolution (LTE) advanced standards. The 3GPP is responsible for developing and optimizing future radio access methods (RATs) and the evolution of the current system. Peak spectrum efficiency targets in downlink and uplink transmission for LTE Advanced systems were set at 30bps/Hz and 15bps/Hz, respectively. Apart from multiple access schemes and enhanced multiple-input multiple-output (MIMO) channel transmission techniques, extensive coordination between numerous cell sites known as coordinated multipoint transmission and reception were accepted as the fundamental techniques for LTE. The medium access mechanism in LTE-A adopts OFDMA for downlink and SC-FDMA for uplink. [6]

2. Fifth-Generation (5G)

5th generation mobile network is currently the most recent advancement in cellular networks and denotes the next primary phase of mobile telecommunications standards beyond the upcoming 4G standards. In other words, It is sufficient to state that the introduction of 5G will allow unlimited access to anybody and anything, everywhere, at any time. [7]

This technology has successfully created end-to-end connectivity between regular use things such as smartphones, fridges, LED lights, Cars, Meters, and many more as a unified air interface.

Two influential organizations worked together to standardize 5G. First, ITU-R (International Telecommunication Union– Radiocommunication) developed 5G requirements. It created the requirement for 3G as IMT 2000 in the year 2000 and for 4G as IMT Advance in 2008. Following these requirements, the second organization still playing a significant role in developing 5G technical specifications is carried out by 3GPP (3rd Generation Partnership Project), a consortium of major operators and telecommunication companies. The key factor driving this evolution from the first generation to the fifth generation is the need for enhanced data rate and efficiency. The following figure highlights the minimal technical requirements for IMT2020.

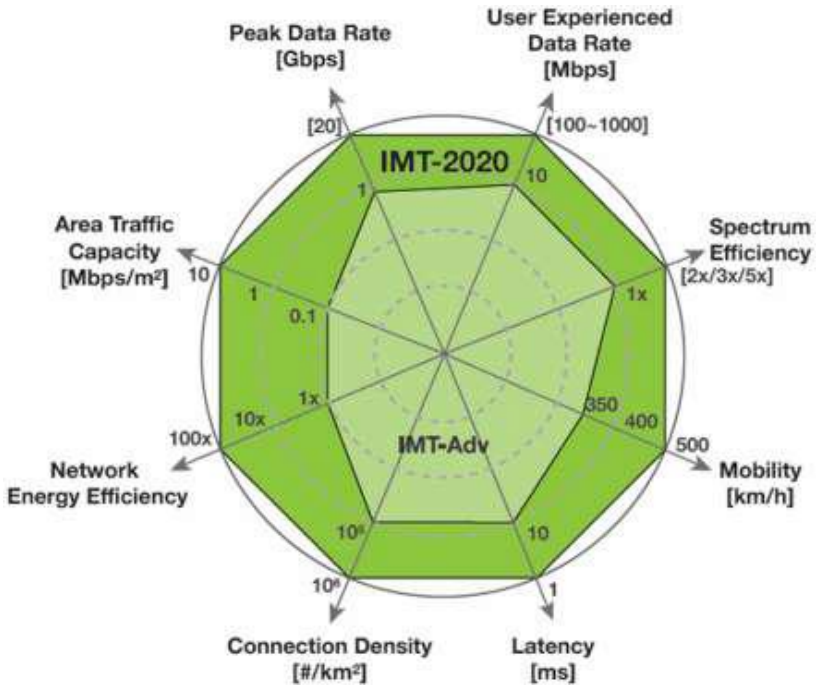


Fig. 9. ITU-R minimal technical requirements for IMT 2020 [8]

Now we see that the peak data rate is 20 gigabits per second, but the actual user data rate varies between hundred and 1000 megabits per second. In the case of 5G in a cell in one meter per square, a user must experience a minimum of 10 megabits per second. Similarly, network efficiency in terms of power-saving should be a hundred times higher than LTE Advanced, especially in the case of IoT, for it to survive longer on battery power. The 5G system must support one million devices per square kilometer. Latency should not exceed one millisecond from the mobile device to the base station. The network should provide service to those mobiles moving up to a speed of 500 kilometers per hour. Regarding spectrum efficiency, LTE-Advance should have three times the efficiency. The better a system's spectrum efficiency, the higher the data rate it can handle.

2.1 5G Usage Scenarios

2.1.1 Enhanced Mobile Broadband (eMBB)

- Expected throughput of 5 Gbps +
- UHD/3D video including broadcast services
- Virtual Reality
- Augmented Reality
- Tactile Internet
- Cloud gaming
- Broadband kiosks
- Vehicular

2.1.2 Ultra-Reliable and Low Latency Communications (uRLLC)

- Industrial control
- Remote manipulation
- Emergency Applications such as e-health, hazardous environments, rescue missions, etc.
- Self-driving vehicles

2.1.3 Massive Machine Type Communications (mMTC)

- A large number of connected devices
- Long batter life
- High connection density
- Agricultural, Subway and Stadium Service, Wearables, Inventory Control, etc

Use cases linked to various scenarios

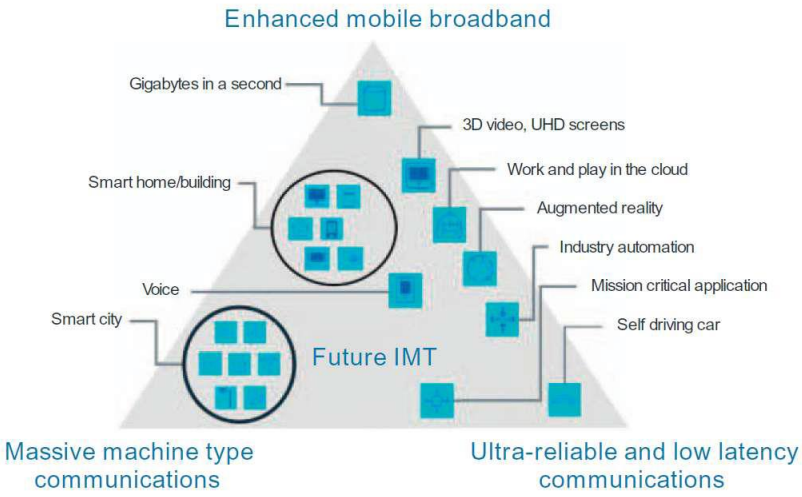


Fig. 10. 5G usage scenarios [8]

There can be secondary use cases based on these three primary use cases. For example, when it comes to augmented reality, we need enhanced mobile broadband, higher data flow, and a high level of reliability and minimal latency.

2.2 Introduction To 5G Architecture

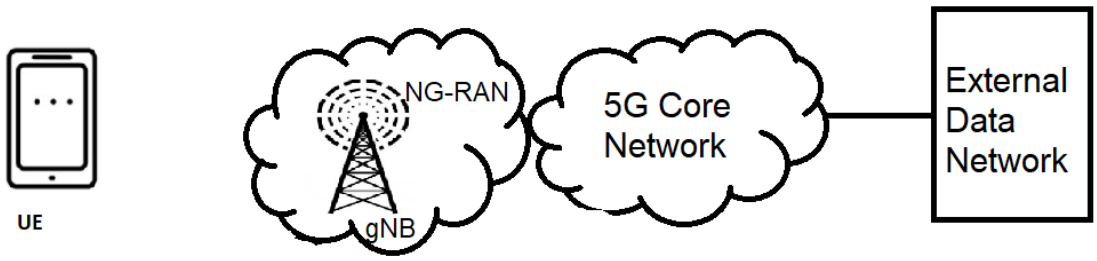


Fig. 11. 5G General Architecture

2.2.1 5G Architecture Main Components

1) UE (User Equipment)

2) NG-RAN (Next Generation-Radio Access Network)

- Uses 5G NR (New Radio) Technology
- Based on gNB (Next-generation NodeB)

3) 5GC (5G Core Network)

- All IP Architecture

2.3 5G Deployment - Non-Standalone (NSA) And Stand Alone (SA) Modes

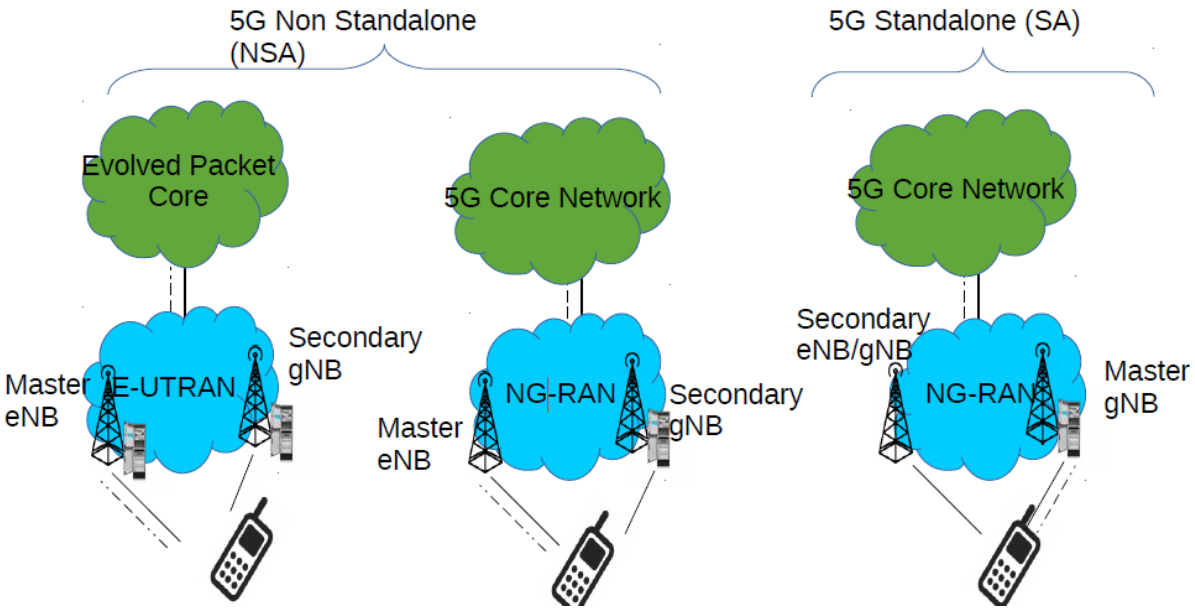


Fig. 12. Non-Standalone (NSA) and Stand Alone (SA) modes.

2.3.1 Non-Standalone Mode

In this scenario, the evolved packet core is operational, which is the core of 4G networks, and prefers the access network of the 4G LTE, the E-UTRAN (the access network of the 4G). The 4G eNodeB is the primary node, while the gNB (5G next-generation NodeB) is the secondary node. Therefore, the UE receives the control signaling and data from the primary node eNB but only data from the secondary node gNB. The primary node also has the information on what type of service and the QoS the UE will receive. In the other scenario, the 5G core is operational, and the access network has been upgraded to the NG-RAN. Hence, these are called the non-standalone because the 4G eNodeB serves the primary component, and the 5G gNodeB plays the secondary component. The NSA architecture is very attractive to service providers and commercial mobile network operators as it requires minimal or no modification to the EPC.

2.3.2 Standalone Mode

The 5G core is operational in the other scenario, and the access network is a 5G NG-RAN. The primary node in this scenario is a gNB, whereas the secondary node might be a gNB or a 4G eNB. As a result, this is a standalone architecture since the 5G gNB is the primary component and the 4G LTE eNB or 5G gNB is the secondary component. This concept where mobile is receiving the data from these NodeB's is very much related to dual connectivity. The SA mode, which uses the NG core, is the long-term goal for 5G deployment and one of the reasons for network operators to migrate from the EPC to NG core, as it will offer high mobility, low latency, and ultra-reliable services.

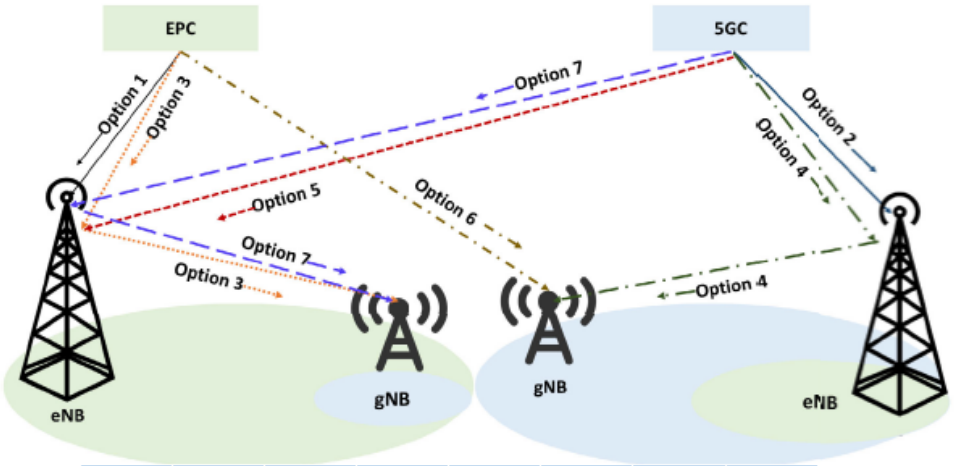
2.4 Key Features Of 5G NG-RAN.

2.4.1 Dual Connectivity

The 5G cell in the early stage of its deployment would not spread lavishly and leave coverage holes to be covered. The existing 4G Long-Term Evolution (LTE) network can internetwork with a 5G network, and the operators can provide seamless services to their users. Fully deployed LTE and 5G compatibility will give network operators quick and seamless coverage and economic viability. [9]

In dual connectivity, one of the NodeB's acts as the master, and the other act as the secondary NodeB that allows mobile devices to utilize both 5G and 4G coverage.
The Master NodeB determines a UE's data and QoS, and bearer splitting can occur at either NodeB's.

2.4.2 Deployment Options



	Option1	Option2	Option3	Option4	Option5	Option6	Option7
Core	EPC	5GC	EPC	5GC	5GC	EPC	5GC
MN	eNB	gNB	eNB	gNB	eNB	gNB	eNB
SN	NA	NA	gNB	eNB	NA	NA	gNB

NA: Not applicable

Fig. 13. 5G deployment options [9]

2.4.3 Small Cells

These cells are cellular radio access codes with low power ranging from 10 meters to a few kilometers and operate in licensed and unlicensed spectrum. Moreover, they have the 5G air interfaces, i.e., they can provide high data rates and low bit error rates as channel impairments of the wireless channel are less due to a small radius of cells.

2.4.4 Increased Spectrum

5G NR can operate in two frequency ranges
1) FR1 450-6000MHz
2) FR2 24250-52600MHz

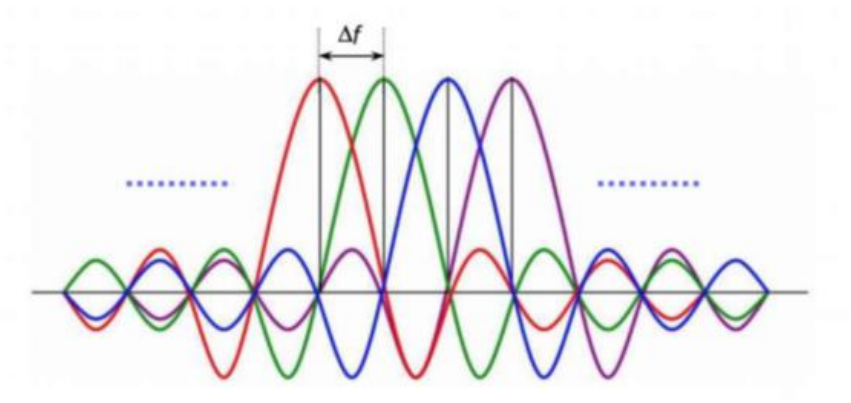
The challenges associated with these frequency ranges are described below :

Below 1 GHz	Excellent in-building penetration, wide area coverage tens of kilometers, limited spectrum availability
1-6 GHz	Adequate coverage and spectrum availability
6-100 GHz	Low range (hundreds of meters), abundant bandwidth hence increased data rate

2.4.5 OFDMA And Flexible Numerology In 5G

Orthogonal Frequency Division Multiplexing (OFDM) is an efficient modulation technique in 5G NR, the access technology used between the gNB and the user equipment on the air interfaces. The frequency band that is in a cell is divided into orthogonal subcarriers. If one subcarrier is at its peak, the other subcarriers are null. As known, 5G has a wide frequency range. However, due to doppler shift and phase error (difference in the phase of the oscillators of the transmitter and the receiver), there is an increase in inter-frequency interference as the frequency rises.

One possible solution to this problem is to use scalable carrier spacing, or in other words, flexible numerical operations characterized by the factor μ. The cyclic prefix is the guard period between two symbols. Its purpose is to ensure that one symbol does not overlap with the subsequent symbol, especially in the multipath environment.



μ	$\Delta f = 2^{\mu} \cdot 15$ [kHz]	Cyclic prefix
0	15	Normal
1	30	Normal
2	60	Normal, Extended
3	120	Normal
4	240	Normal

Table 1. Supported transmission numerologies [10]

2.4.6 Frame Structure

Frames in 5G have a longer duration of 10 ms for both downlink and uplink transmission. Each frame consists of 10 ten subframes of 1 ms duration. One set of frames for the uplink (0-4) and one for the downlink (5-9) on a carrier.

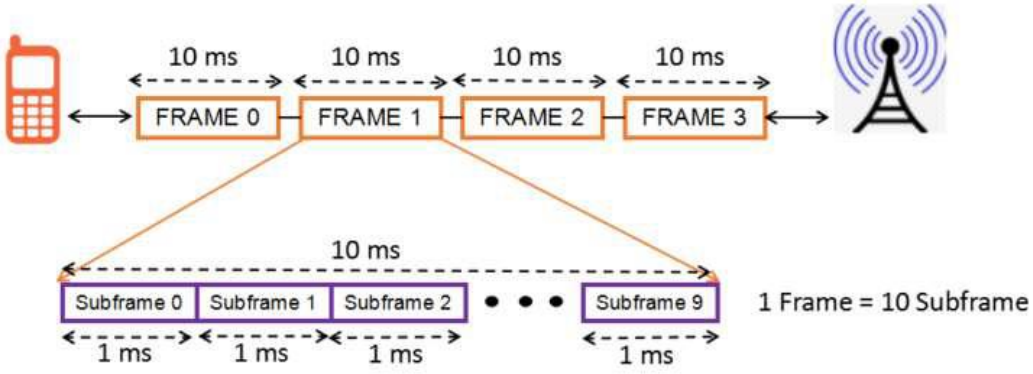


Fig. 14. Frame Structure 5G NR. [11]

2.4.7 Slots

Inside a subframe, slots are numbered in increasing order for subcarrier spacing configurations within one frame. There are 14 OFDM symbols per slot (Normal Cyclic Prefix). Generally, Slot length gets shorter as subcarrier spacing gets wider.

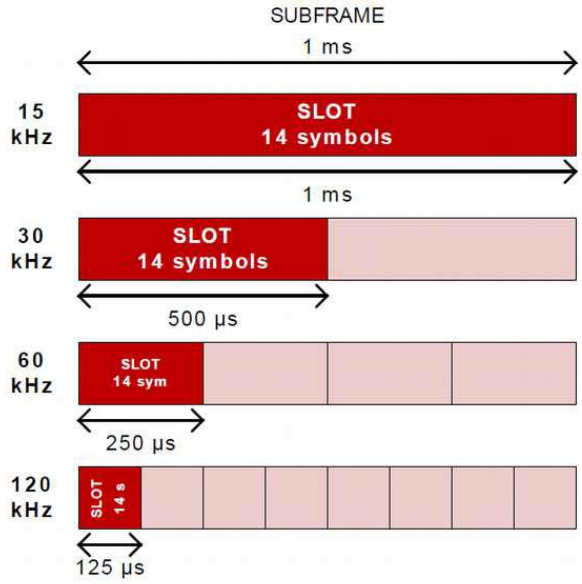


Fig. 15. Relationship between slot and subframe

μ	N_{symb}^{slot}	$N_{slot}^{frame,\mu}$	$N_{slot}^{subframe,\mu}$
0	14	10	1
1	14	20	2
2	14	40	4
3	14	80	8
4	14	160	16

Table 2. Number of OFDM symbols per slot [10]

2.4.8 Resource Block

In the frequency domain, a resource block comprises 12 consecutive subcarriers. The resource block in 5G New Radio is only defined for the frequency domain because the duration of a time slot diminishes as the subcarrier spacing increases in the time domain.

2.4.9 Resource Element

A resource element is a resource grid and subcarrier spacing configuration component. In other words, one frequency domain subcarrier and one-time domain OFDM symbol develop the smallest unit.

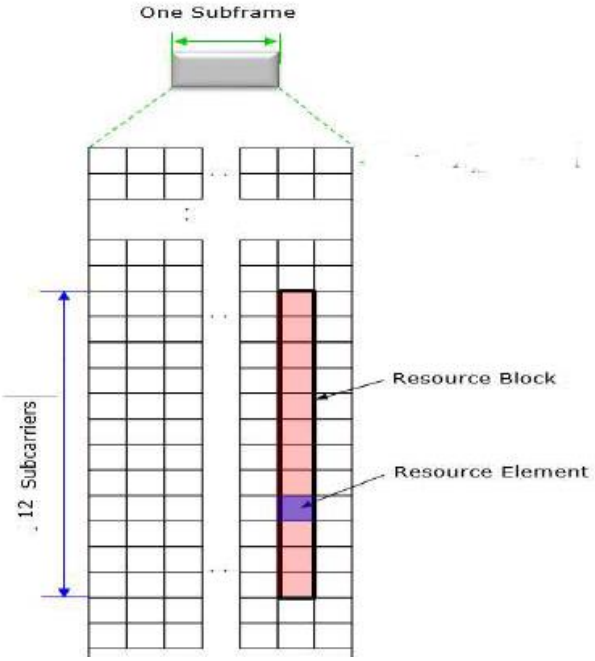


Fig. 16. Resource element

2.4.10 Modulation

Modulation refers to the number of bits transmitted in a particular resource element, whether data or parity bits. 5G new radio supports QPSK, (16-64-256) QAM modulation techniques. For example, 5G may employ up to 256 QAM (8 bits/symbol) under ideal wireless channel conditions and lower modulation as the link deteriorates.

2.4.11 Massive MIMO And Beam Forming

The key technology for 5G is Massive MIMO (Multiple Input Multiple Output). A large number of antennas that are installed on the base station_work together to improve both coverages and increase the data rate. Massive MIMO and beamforming are terms that are often used together in 5G.

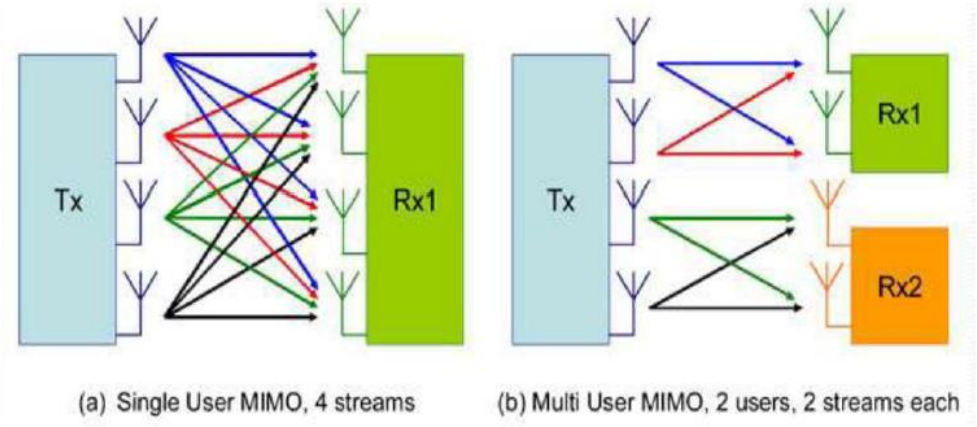


Fig. 17. SU-MIMO and MU-MIMO

SU-MIMO vs. MU-MIMO

In Single User MIMO, both the transceiver and receiver have multiple antennas. Multiple data streams are transmitted simultaneously using the same resources (time/frequency) to double or quadruple the peak throughput.

In MU-MIMO, multiple data streams are transmitted simultaneously by a base station using the same resources (time/frequency), one per receiver. Hence there is an increase in the total throughput.

MIMO techniques are suitable for wireless channels with rich multipath and scattering, increasing the system's performance. MIMO technology is affected by the change in frequency used for transmission. When the frequency is below 6 GHz, many multi-paths are produced, resulting in increased scattering; therefore, 8x8 MIMO can be achieved. Whereas in the case where the frequency is above 6 GHz, there is reduced scattering, and hence a smaller number of multi-paths are produced, leading to a maximum of 2x2 MIMO.

MIMO and beam forming are not exclusive to one another as some antennas can be used for beamforming and others for MIMO. Massive MIMO technology uses a large antenna array on the base station, typically comprising 16, 32, or 64 or more array elements. The number of antennas array elements is much larger than the number of UEs. As a result, narrower beams can be directed at different UEs, increasing the SNR ratio and causing less interference.

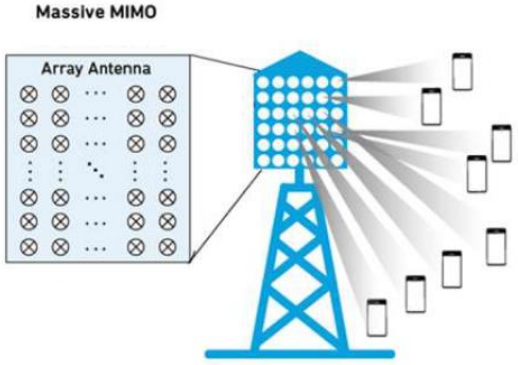


Fig. 18. Antenna Array

2.5 5G Reference Network Architecture

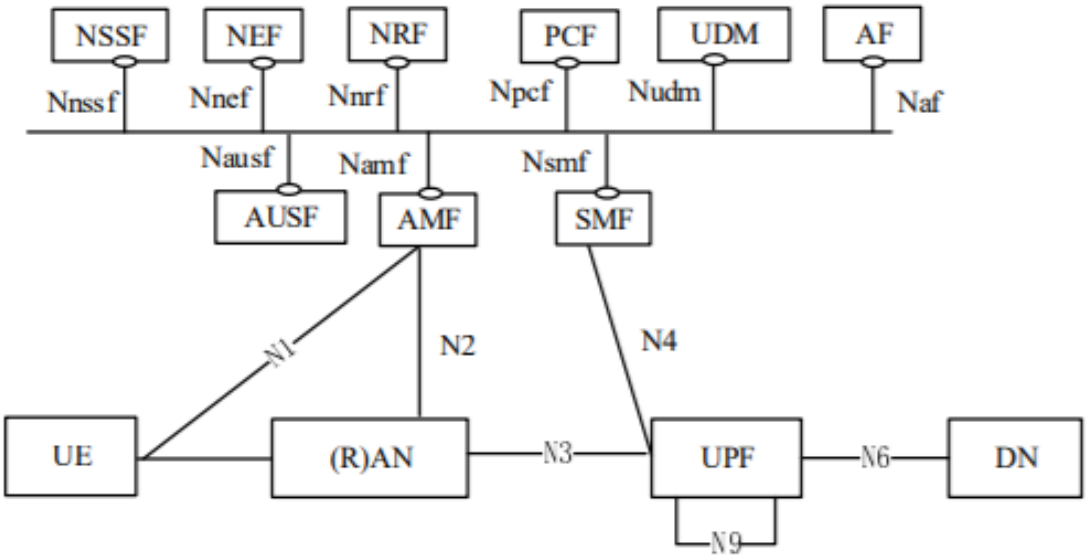


Fig. 19. 5G Reference Network Architecture [12]

Network nodes and their functions [12] :

2.5.1 Next-Gen Node Basestation (gNB)

Following functionalities :

- Radio Transmission/Reception
- Digital Signal Processing
- Encryption and data compression
- Process Access Stratum Signaling
- Relay Non-Access Stratum signaling to Core
- Radio Resource Management
- Communication with Core Network and nearby base stations

2.5.2 Core Access and Mobility Management Function (AMF)

Following functionalities :

- Mobility Management
- Registration management.
- Connection management.
- Reachability management.
- Termination of RAN control plane interface (N2) and NAS (N1), NAS ciphering, and integrity protection.
- Lawful intercept
- Access Authentication/Authorization

2.5.3 User Plane Function (UPF)

Following functionalities :

- Packet routing & forwarding.
- Packet inspection.
- Policy rule enforcement.
- Lawful intercept (User Plane).
- Traffic reporting and accounting
- QoS management for the user plane.
- Anchor point for Intra/Inter-RAT mobility.
- Transport level packet marking in the uplink/downlink.

2.5.4 Session Management Control Function (SMF)

Following functionalities :

- Session Management
- Allocation of IP for UE & management
- Control and Selection of User Plane function
- Termination of interfaces (Policy control)
- Lawful intercept
- Termination of Session Management
- Downlink Data Notification
- Roaming functionality
- Charging data collection
- Support of charging interface

2.5.5 Data Network (DN)

Provides operator services, Internet access, or other services.

2.5.6 Authentication Server Function (AUSF)

Perform the UE authentication process.

2.5.7 Unified Data Management (UDM)

Following functionalities :

- Access authorization based on subscription data
- De-concealment of privacy-protected subscription identifier (SUCI).
- Subscription and SMS management.
- Generation of 3GPP AKA Authentication Credentials.
- Include Authentication Credential Repository and Processing Function (ARPF).

2.5.8 Policy Control Function (PCF)

Following functionalities :

- Accesses subscription information
- Policy rules to control plane function to enforce them

2.5.9 Application Function (AF)

Following functionalities :

- Requests dynamic policies and charging control
- Communication with the core network to request a packet flow
- IMS Node requesting voice call

2.5.10 NRF (Network Repository Function)

Following functionalities :

- Maintains profiles for Network Functions
- Receive Network Function Discovery Request
- Location and Identification of stored data/information.

2.5.11 NSSF (Network Slice Selection Function)

Following functionalities :

- Selection of set for Network Slice instances serving the UE
- Determining the Allowed/Configured NSSAI

2.5.12 NEF (Network Exposure Function)

Following functionalities :

- Exposes the capabilities of 5G core network functions to an external AF.
- Acts as the middleman for information exchanged between 5GC and AF.
- The received information is stored as structured data

2.5.13 Service-based interfaces

The system architecture includes the following Service-based interfaces:

Namf: Service-based interface exhibited by AMF.

Nsmf: Service-based interface exhibited by SMF.

Nnef: Service-based interface exhibited by NEF.

Npcf: Service-based interface exhibited by PCF.

Nudm: Service-based interface exhibited by UDM.

Naf: Service-based interface exhibited by AF.

Nnrf: Service-based interface exhibited by NRF.

Nnssf: Service-based interface exhibited by NSSF.

Nausf: Service-based interface exhibited by AUSF.

2.5.14 Reference points

The System Architecture contains the following reference points:

N1: A Reference point between the UE and the AMF.

N2: A Reference point between the RAN and the AMF.

N3: A Reference point between the RAN and the UPF.

N4: A Reference point between the SMF and the UPF.

N6: A Reference point between the UPF and a Data Network.

2.6 PDU (Protocol Data Unit) Sessions

As already mentioned, 5G is an all IP system that provides the services like voice, data, and multi-media are carried over the IP packets. The network must establish PDU sessions for the user to provide those services, and each PDU session may consist of one or more quality of service flows. Quality of service ID identifies each quality-of-service flow. QoS is characterized by data rate (guaranteed or non-guaranteed), latency, and priority.

Fig. 20. PDU establishment and QoS flow

2.7 Control And User Plane Separation

2.7.1 Control Plane Function

- Authentication
- Connection management
- QoS management
- Mobility management

2.7.2 User Plane Functions

- data traffic forwarding

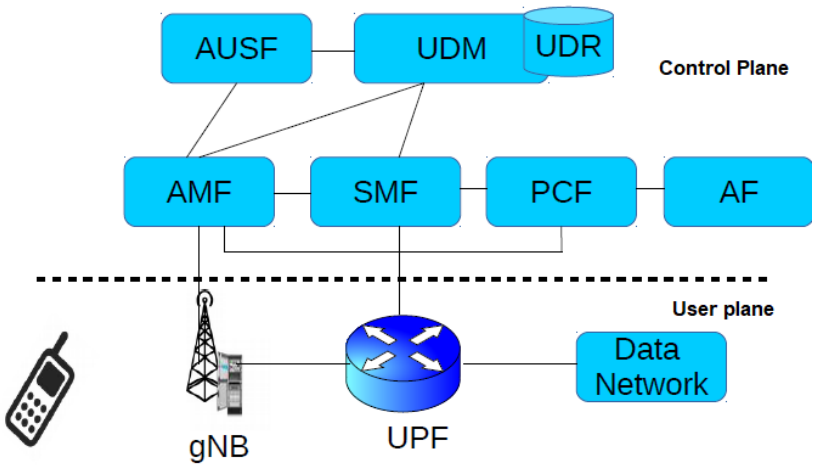


Fig. 21. User Plane (UP) and Control Plane (CP) separation.

Separating the User Plane functions from the Control Plane functions allows

- Independent scalability
- Evolution
- Flexible deployments.

2.8 Identifiers In 5G

UE identifier

UE comprises 2 elements :

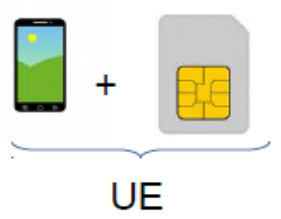


Fig. 22. UE elements

- Mobile device
- SIM (Subscriber identity module)

The mobile identity by which the network uniquely identifies the mobile set or device is known as Permanent Equipment Identity (PEI), which supports the IMEI format. Similarly, the network's identity identifying the subscriber using SIM is SUPI.

SUPI can be IMSI for the 3GPP access or NAI (Network Access Identifier) for non-3GPP Access.

When a mobile device is powered up and tries to connect to a network, an IMSI (International Mobile Subscriber Identity) is generated in the previous generation. However, this IMSI could be traced using an IMSI catcher for attack purposes.

To solve this problem in 5G, when the mobile-first registers to the network, its SUPI is encrypted using a secret key within the SIM called SUCI (Subscription Concealed identifier).

When the core network registration is complete, AMF assigns a 5G S Temporary Mobile Subscriber Identity (5G-S-TMSI) unique inside an AMF region and varies with time. If the AMF area of mobile is unknown, then 5G Global Unique Temporary Identifier (5G-GUTI) is assigned by AMF.

2.9 Tracking Areas

The service area of a Mobile is divided into the tracking areas, which do not overlap. A tracking area at the minimum may contain one cell and, at the maximum many.

Mobile equipment can be in one of two states.

- Ideal mode
- Connected mode

In the case of ideal mode, the location of mobile equipment is known to the level of the tracking area and not the cell. The purpose of the tracking area is to reduce the network's signaling load during an incoming call when the network must page all of the mobiles in a cell. Paging message would contain 5G-S-TMSI of mobile for which the call is generated, and mobile will then respond to the network to identify the location. Tracking Area is updated when a mobile enters a cell with a different location.

The core network can list the specific UE tracking area to roam without updates for mobile equipment moving at high speed.

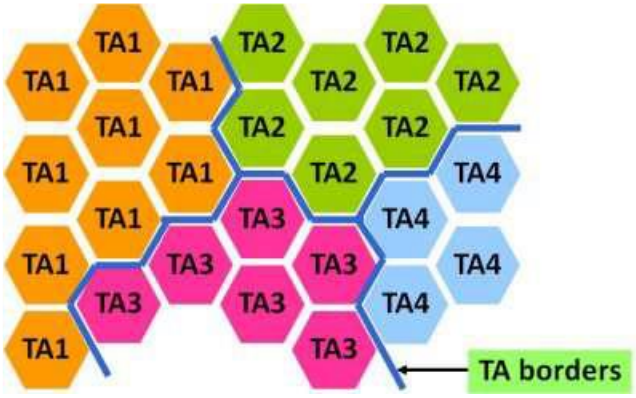


Fig. 23. Tracking areas in 5G [13]

2.10 Network Identifiers

A mobile network comprises two components, a radio access network and a core network. The identity of this mobile network is called the PLMN-ID (Public Land Mobile Network Identity). It combines the mobile country code and the mobile network assigned to this operator.

Public Land Mobile Network Identity (PLMN-ID) :

(PLMN-ID)=MCC|MNC

AMF Identifier (AMI) :

Globally Unique AMF Identifier

(GUAMI)=MCC|MNC|AMI

Tracking Area Identity (TAI):

Each tracking area has a code called the tracking area code.

(TAI)=MCC|MNC|TAC

2.11 5G Network Slicing

Network Slicing is considered one of the key features by 3GPP in 5G. Network Slice is a separate network (logical) serving a defined customer. It consists of all the network resources required, configured together. A UE may access multiple slices over the same gNodeB.

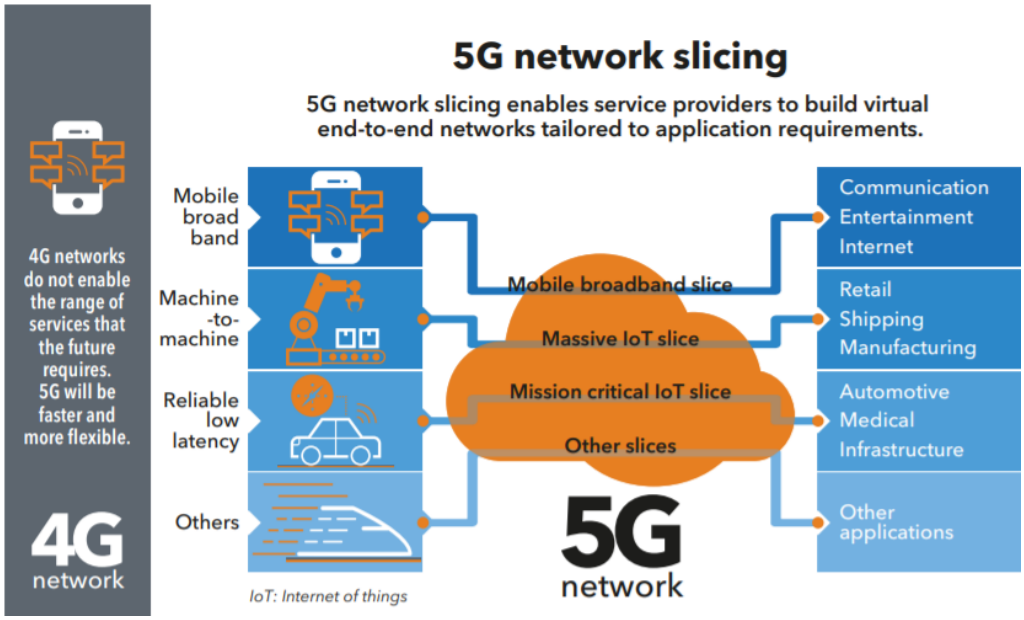


Fig. 24. Network Slicing [14]

2.11.1 Features

- Complete network within a provider
- Resources optimized for one use case
- Isolated but may share resources
- User experience is as a separate network
- On-demand allocation of resources

According to [12], A Network Slice is defined within a PLMN and should include:

- 5G Core
- 5G RAN Network Control Plane and Network Control Plane and User Plane Network

3GPP release 15 [12] defines Network Function, Slice, and Slice Instance as

- **Network Function**: A 3GPP-defined processing function, defined functional behavior, and 3GPP-defined interfaces are all examples of 3GPP-defined processing functions. A network function can also be implemented as a dedicated network element, a software instance running on dedicated hardware, or a virtualized function instantiated on an appropriate platform, such as cloud infrastructure.
- **Network Slice**: A logical network with specific network characteristics and attributes.
- **Network Slice instance**: A collection of network function instances and required resources, such as compute, storage, and network resources, that are part of the deployed network.

2.11.2 Identification And Selection Of A Network Slice

Single Network Slice Selection Assistance Information (S-NSSAI) identifies a Network Slice.

An S-NSSAI is comprised of:

2.11.2.1 A Slice/Service type (SST)

Provides information on Slice's behavior in terms of features and services.

2.11.2.2 A Slice Differentiator (SD)

Optional information to differentiate multiple Network Slices.

Fig. 25. Single Network Slice Selection Assistance Information

An S-NSSAI can have standard values (only SST and no SD) or non-standard values (either SST alone (Non-standard) or SST + SD)

SST Value	Service Type
1	eMBB
2	URLLC
3	MIOT
4	V2X
5-127 Reserved	
128-255 Operator Specific	

Table 3. SST values and Service types

*3GPP has reserved values 0 to 127 for standardized SST [12]

2.11.3 Network Slice Subnet Instance (NSSI)

An NSI comprises one or multiple Network Slice Subnet Instances (NSSIs). An NSSI may

- Consist of Network Functions and other Network Slice Subnet Instances
- Be shared by two or more Network Slice instances
- Contain Core Network functions or Access Network functions, or both

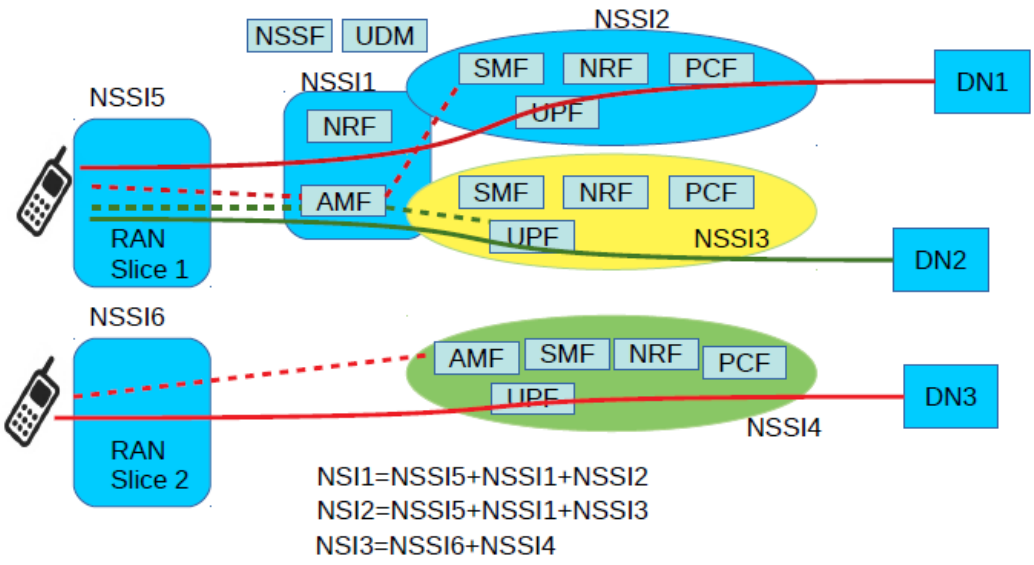


Fig. 26. Network Slice Subnet Instance (NSSI) example

A maximum of 8 slices can be assigned to a single UE at a time. UE must support PDU sessions associated with these slices. A general AMF instance supports all slices assigned to a UE and can have separate SMF/UPF instances. A PDU Session is associated with only one S-NSSAI and one DN (Data Network).

3. Security In 5G Networks

The 5G System's security characteristics include [12]:

- UE authentication by the network
- Generation and distribution of security context.
- User Plane Integrity and confidentiality protection.
- Integrity protection and signaling confidentiality of Control Plane.
- User identity confidentiality

Non-Access-Stratum (NAS) security is initiated to secure multiple instances when a UE is connected to an NG-RAN. Each instance is created based on the security context of that corresponding SEAF obtained from UE authentication. E.g., UE served by common AMF [10].

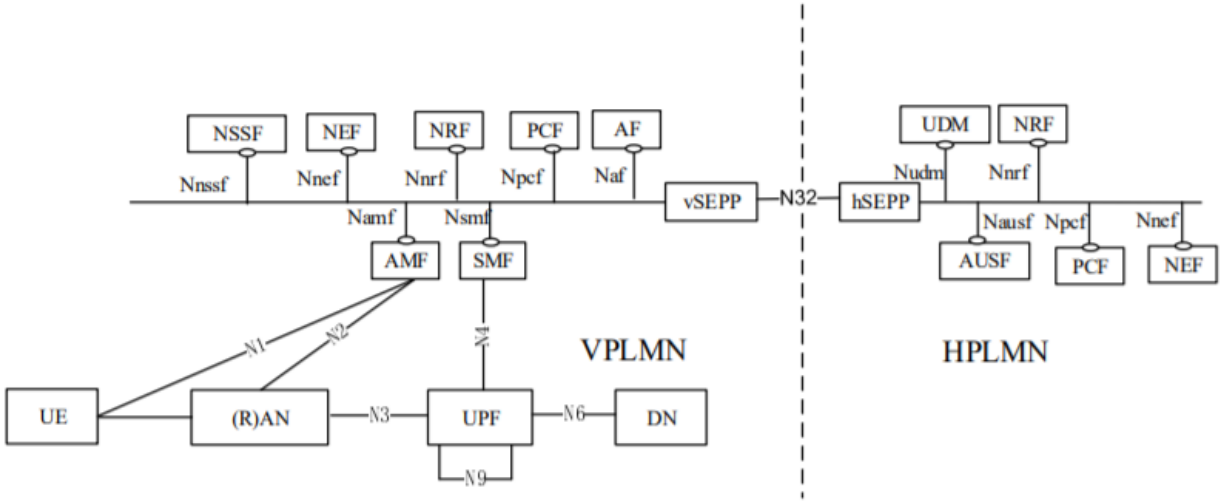


Fig. 27. 5G Roaming Architecture [12]

3.1 Logical Entities For Network Access Security

Developing independent logical security entities ensures a logical security architecture within the 5G Core Network Functions.

Logical Entities :

- ARPF (Authentication credential Repository and Processing Function)
- AUSF (Authentication Server Function)
- SEAF (Security Anchor Function)
- SIDF (Subscription Identifier De-concealing Function)

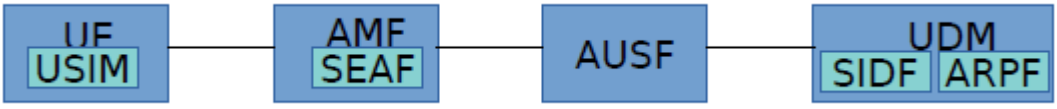


Fig. 28. 5G Security Logical Entities

3.1.1 ARPF (Authentication Credential Repository and Processing Function)

Subscriber's security credentials are stored in ARPF. Subscriber credentials consist of at least one long-term key or master key and the subscription identifier SUPI in USIM and ARPF. This key authenticates UE and the 5G core network and uniquely identifies a subscription. [15].

ARPF services are provided via the UDM, but no open interface is defined between UDM and ARPF. Instead, a master key is used to maintain confidentiality, encryption, decryption, authentication, and integrity protection.

3.1.2 SIDF (Subscription Identifier De-Concealing Function)

Subscription Identifier De-concealing Function (SIDF) is a service provided by the network function UDM, which is responsible for de-concealing the SUPI from the SUCI in the local network of the subscriber [15].

The SIDF is responsible for the following requirements :

- A service offered by UDM for de-concealment of the SUCI.
- Based on the protection used to generate SUCI, SIDF shall resolve the SUPI from the SUCI.

*SUPI is not sent over the air interface. Instead, it is encrypted as the SUCI and SIDF are responsible for de-conceal the SUCI.

3.1.3 AUSF (Authentication Server Function)

It's a standalone network function in the 5G architecture that handles authentication on data received from the UE and UDM, and ARPF in the local network.

Requirements on AUSF

- Shall handle authentication requests.
- Provide SUPI only after authentication confirmation
- Successful or unsuccessful authentication signaling to UDM

3.1.4 SEAF (Security Anchor Function)

The SEAF manages authentication functions in the serving (visited) network through the AMF based on the UE and AUSF information. SEAF serves as a secure anchor for UE.

SEAF shall fulfill the following requirements:

- SEAF supports primary authentication using SUCI.

3.2 Network Access Security

A set of security features, in particular, to authenticate and access services via the network securely enabled on a UE to protect against attacks on interfaces (radio) [15].

Mutual authentication is required for UE and Network to confirm each other's identities. Now, this requires signaling between UE and the core network.

UE and network must support two authentication methods:

- 5G AKA (5G Authentication and Key Agreement) for 3GPP Access
- EAP-AKA (Extensible Authentication Protocol– Authentication and Key Agreement) for non-3GPP Access

Once the authentication procedure is complete, the ciphering and integrity protection keys are generated. As recommended by 3GPP, signal integrity protection is mandatory except in urgent cases, and signals between network functions must be encrypted whenever permitted by regulations.

3.2.1 Initiation Of Authentication

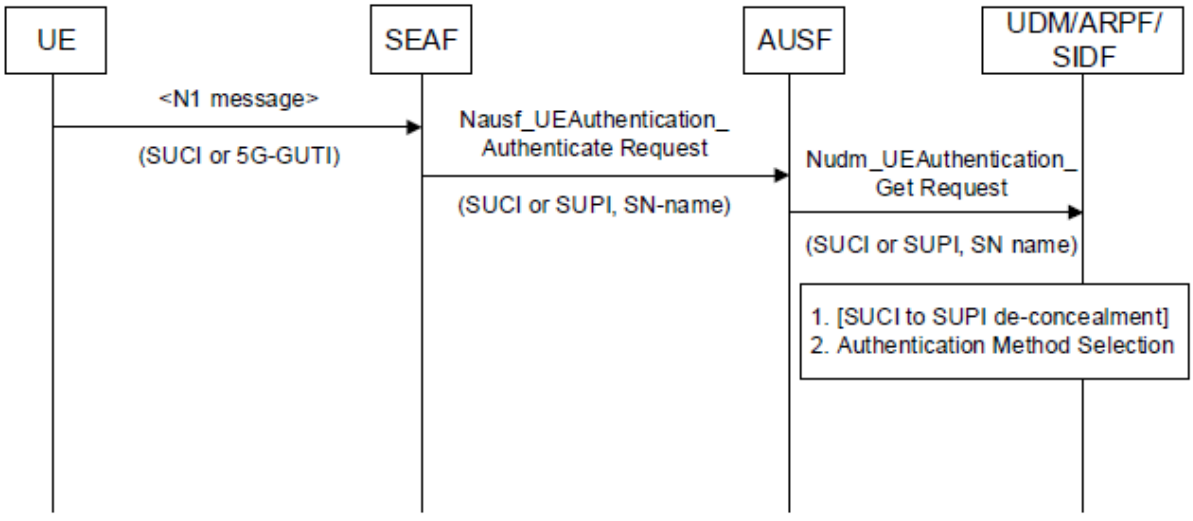


Fig. 29. Initiation of authentication procedure [15]

3.2.2 Concealment and De-concealment of SUPI

- If the UE has a valid 5G-GUTI from previous registration, 5G-GUTI is sent.
- SUPI must be applied if 5G-GUTI is not accessible.
- SUPI is never transmitted in clear text over the radio; instead, an encrypted version known as SUCI is provided.
- SUCI is generated by the UE based on elliptical curve integrated encryption scheme public-key cryptography and a protection scheme.
- The UDM/SIDF can then derive the SUPI from the SUCI using the home network private key.

3.2.3 Authentication Procedure For 5G AKA

The ARPF generates the 5G authentication vector. It then returns a 5G authentication vector, which comprises four parameters [16] :

- A random number challenge to the UE – RAND
- An expected response computed from RAND, K, and the network name - XRES*
- The home network provides an authentication token to verify that the network is genuine - AUTN
- If the UE was identified using its SUCI. Then a starting point for computing the lower level keys – K_AUSF

Then the UDM also returns the corresponding SUPI.

- Uses Milenage functions to derive the following vectors - MAC, XRES, CK, IK, and AK
- UDM uses key Derivation Function HMAC-SHA-256 KDF to derive XRES*
- Key Derivation Function HMAC-SHA-256(K, S) is used by UDM to derive K_AUSF
- RAND (OR) XRES* is input to the SHA-256 algorithm at AUSF to calculate HXRES*. HXRES* is the output of the SHA-256 hash and is 128 bits in length.
- Derivation K_SEAF from K_AUSF requires AUSF to pass the value K= K_AUSF and S = 0x6C || Network Name || Length of Network Name to KDF function.
- XMAC, RES, CK, IK are again calculated by UE using the Milenage functions.
- The Keys calculated above derive RES* by utilizing the HMAC-SHA-256 KDF function by UE and forwarded to AMF.
- Calculation of HRES* from RES* is done by AMF through passing RAND || RES* as input to the SHA-256 algorithm. HRES* is the output of the SHA-256 hash and is 128 bits in length.

*Note that the 5G specifications only allow the UDM to return one authentication vector at a time.

Derived Authentication Vectors	Explanation
IK	Integrity Key Generated
CK	Ciphering Key Generated
AK	Anonymity Key Generated
XRES	Response Generated
MAC	Message Authentication Code Generated
AUTN	Authentication Token Generated
RAND	Random Number Generated

Table 4. Derived Authentication vectors

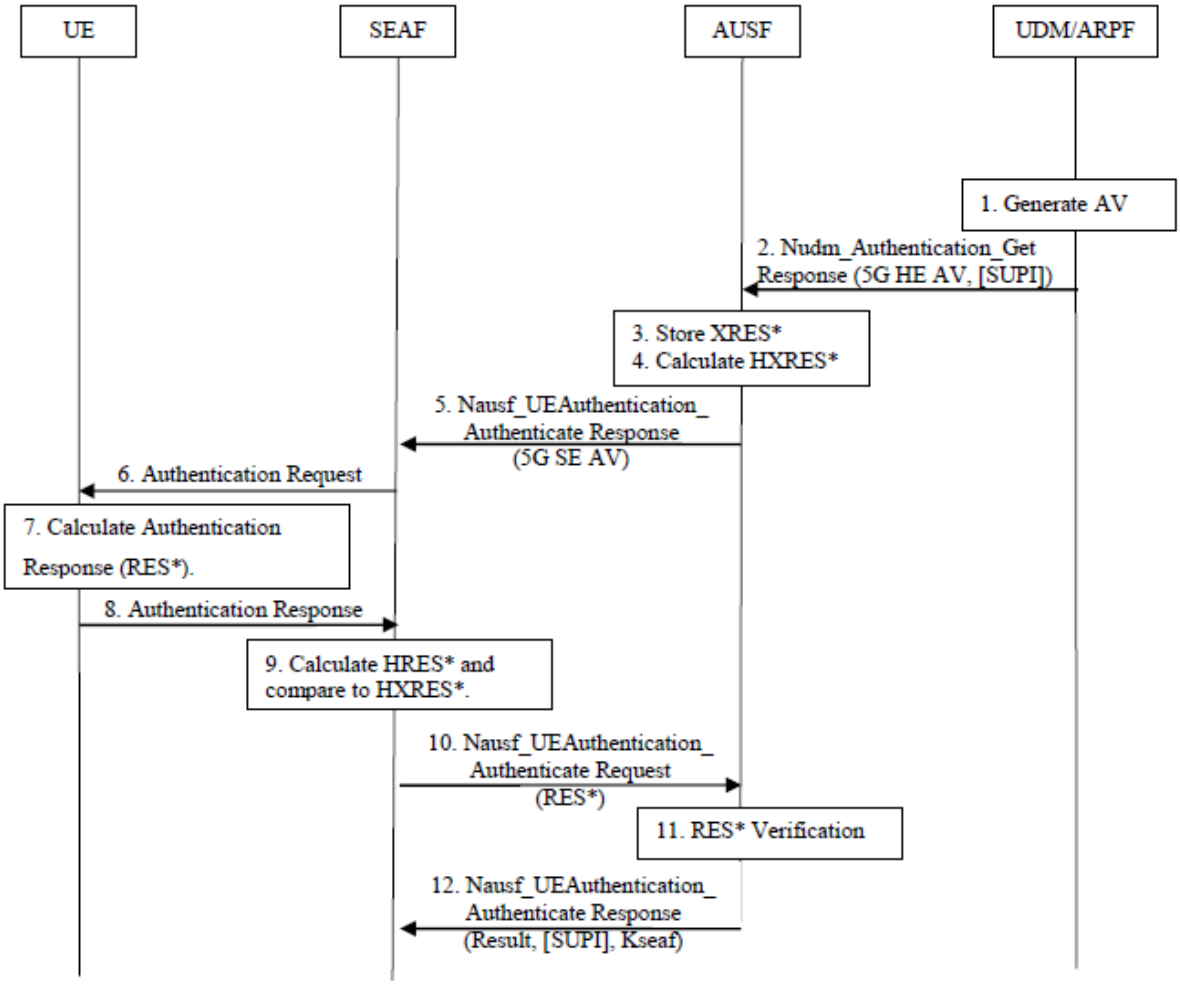


Fig. 30. Authentication procedure for 5G AKA [15]

3.2.4 5G EAP-AKA (Extensible Authentication Protocol – Authentication and Key Agreement) For Non-3GPP Access Architecture

A protocol framework that is designed typically for authentication between the end-user and a network is defined in RFC 3748 by IETF. Primarily introduced for the Point-to-Point Protocol (PPP). EAP is a general authentication framework, not a specific authentication mechanism to implement particular authentication methods. Support for different authentication methods is enabled by EAP and therefore is extensible. In addition, new authentication methods, commonly referred to as EAP methods, can be added.

EAP-AKA is an EAP method for performing authentication based on USIM cards defined by IETF in RFC 5448. Transformed Authentication Vector (AV') is generated by UDM, ARPF and provided to AUSF.

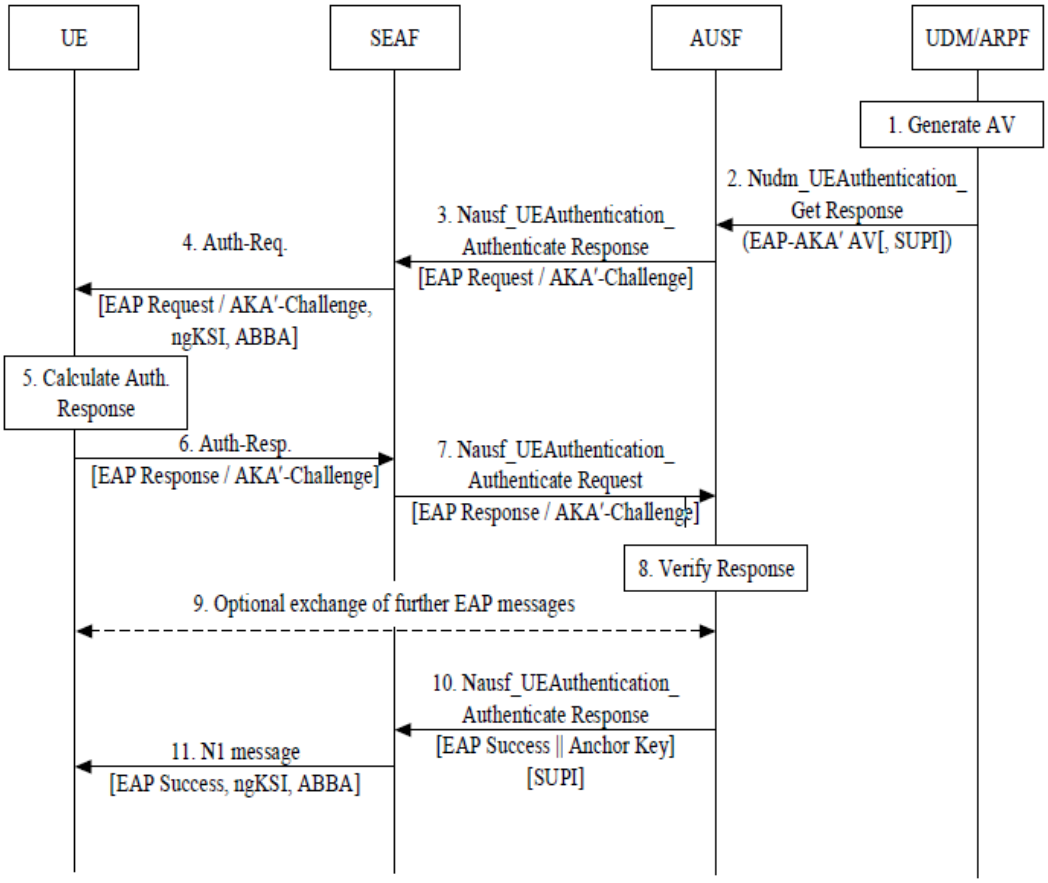


Fig. 31. Authentication procedure for 5G EAP-AKA [15]

3.3 Key Hierarchy

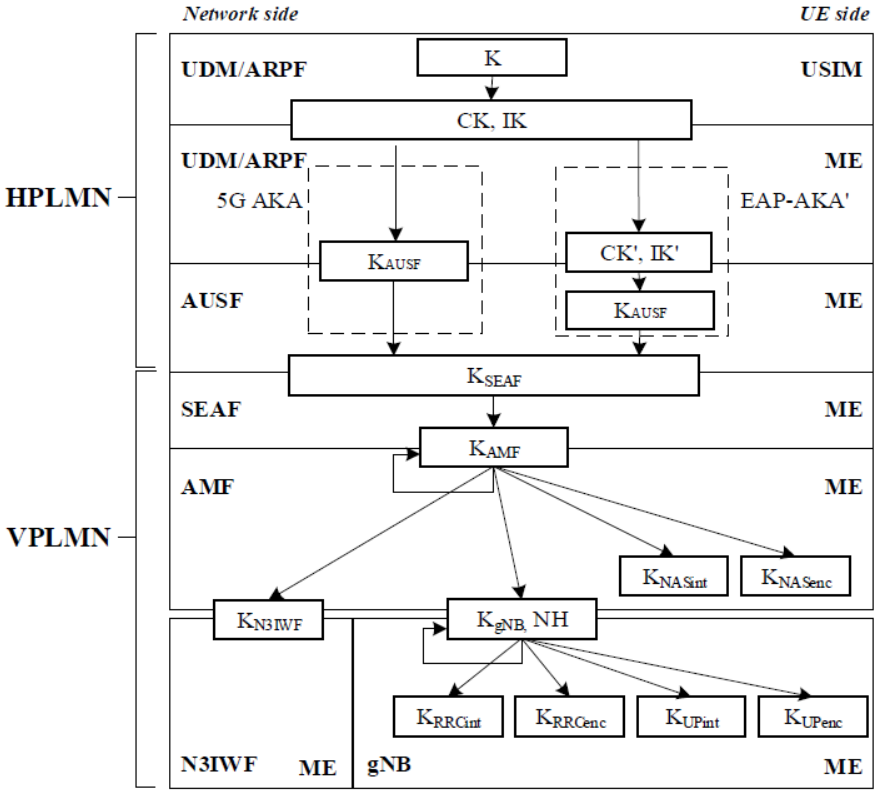


Fig. 32. Key hierarchy generation in 5G [15]

3.3.1 Cryptographic Algorithms and Protection Schemes

The 3GPP standard has specified the following algorithm between UE, RAN, and AMF for confidentiality and integrity of communications [17].

NEA (Encryption Algorithm) - Ciphering algorithm

NIA (Integrity Algorithm) - Integrity algorithm

3.3.1.1 NEA0 / NIA0 (Null Algorithms)

These do not provide encryption or integrity protection. These algorithms should be used in scenarios where it is impossible or unnecessary to authenticate the UE except permits for emergency purposes. The current standards exclude these algorithms for the list as they are weak and add an unnecessary overhead of 32-bits to the MAC header with no security benefits.

3.3.1.2 128-NEA1 / 128-NIA1

Snow 3G stream cipher is used in 128-NEA1/128-NIA1. Snow 3G is a 32-bit word stream encryption that supports 128-bit keys.

3.3.1.3 128-NEA2 / 128-NIA2

128-NEA2/128-NIA2 is based on the Advanced Encryption System block cipher.

3.3.1.4 128-NEA3 / 128-NIA3

ZUC stream cipher generates a 32-bit word used for 128-NEA3/128-NIA3 as it is the least analyzed standard.

3.4 5G NR RADIO PROTOCOL STACK

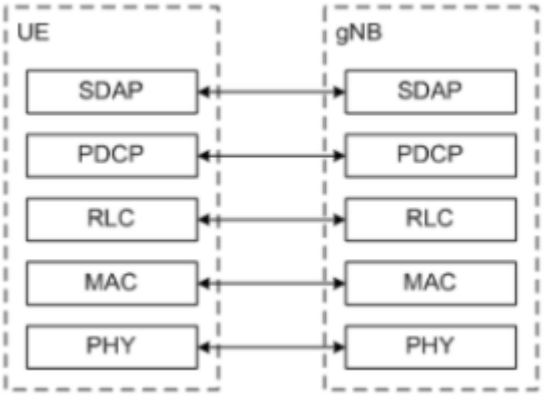


Fig. 33. User Plane Protocol Stack [18]

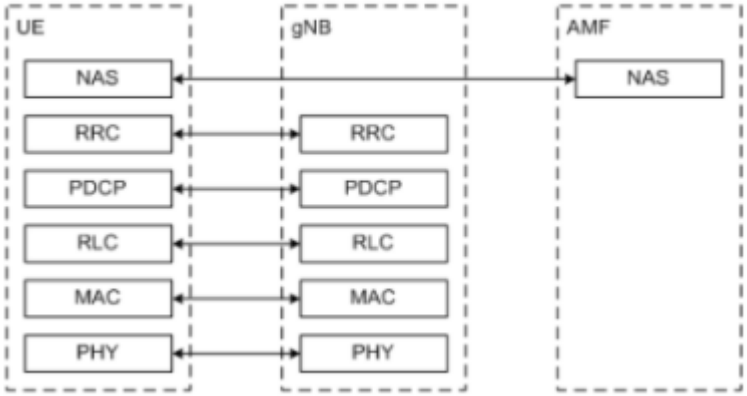


Fig. 34. Control Plane Protocol Stack [18]

3.4.1 RRC Layer 3 Functions

- Access stratum and non-access stratum system information broadcast
- Paging initiation.
- The UE and the RAN connection are established, maintained, and released via radio resource control.
- Creation, modification, and release of E-UTRA and NR dual connectivity
- Key management with security features.
- Setup, configuration, maintenance, and approval of signaling and data radio bearers.
- Mobility functions include handover and context transfers.
- UE cell selection and reselection management.
- Inter-RAT mobility.
- QoS management functions.
- UE measurement details and report control
- Detection and recovery after a radio link failure.
- NAS messaging between NAS and UE.

3.4.2 Layer 2 Functions

Layer 2 is divided into the following sublayers :

- Service Data Adaptation Protocol (SDAP)
- Packet Data Convergence Protocol (PDCP)
- Radio Link Control (RLC)
- and Medium Access Control (MAC)

Figures 35 and 36 describe the Layer 2 architecture for the downlink and the uplink.

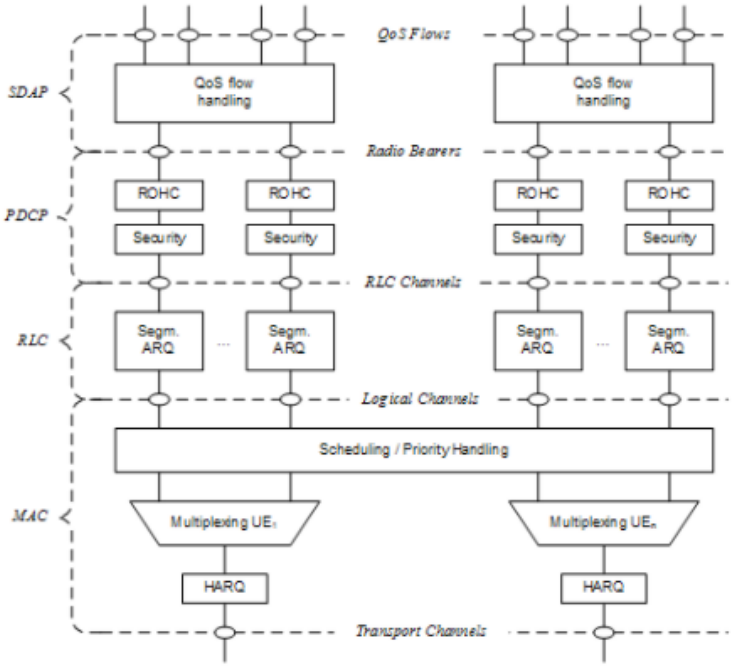


Fig. 35. Downlink Layer 2 Structure [18]

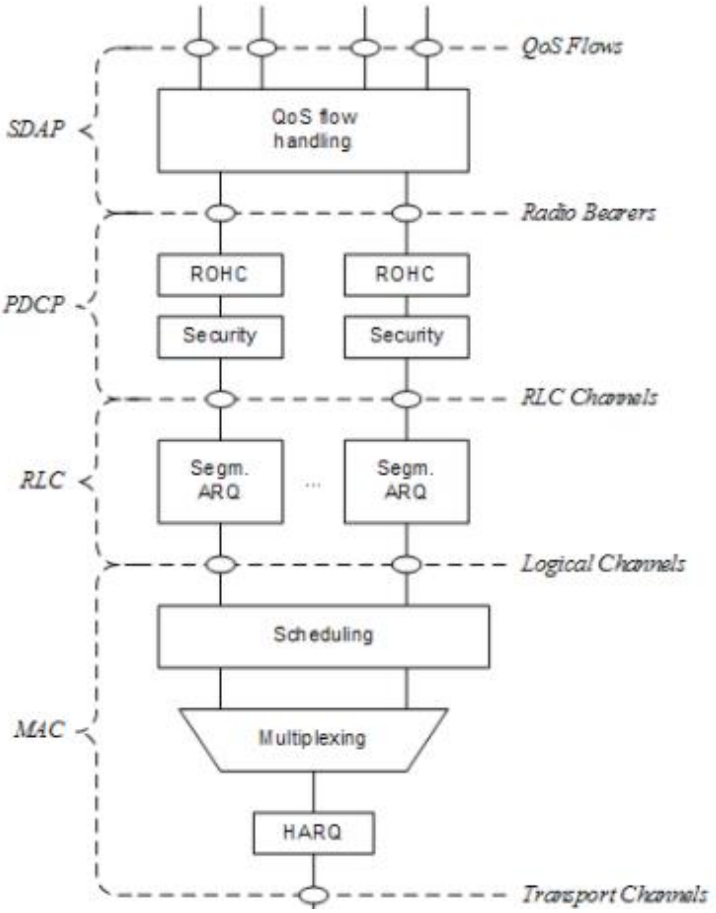


Fig. 36. Uplink Layer 2 Structure [18]

3.4.3 SDAP (Service Data Adaptation Protocol)

SDAP is the topmost sublayer layer in the 5G protocol stack. A single protocol instance of the protocol is configured per single packet data unit session, except for data convergence that can configure two entities.

The main services and functions include:

- QoS flow identification marking
- Ensure correct forwarding
- Mapping of IP flows

3.4.4 PDCP (Packet Data Convergence Protocol)

PDCP is primarily responsible for header compression and security in the protocol stack by appending a sequence number to packet data units.

The user-plane functions in PDCP include:

- Robust Header Compression and decompression.
- User plane data transfer
- Duplicate data detection and reordering based on the sequence number
- Routing of packed data unit
- Assigning sequence numbers to packets
- Retransmission service data unit
- Encryption and decryption
- Discard PDCP service data unit
- Radio link control acknowledgment
- Data recovery
- PDCP packet data unit duplication.

The control plane functions in PDCP includes:

- Encryption and decryption
- Integrity protection.
- Assigning sequence numbers to packets
- Control plane data transfers.
- Detection of duplicate data units.

3.4.5 RLC (Radio Link Control)

RLC is responsible for segmentation packet units and error-checking mechanisms like Automatic Repeat Request (ARQ).

The main RLC functions include:

- Assigning sequence numbers to packets
- Automatic repeat request error correction
- Packet data units transfer for upper layer
- Segmentation and re-segmentation
- Service data unit reassembly
- Discard RLC service data unit
- Recovery and re-establishment

3.4.6 MAC (Media Access Control)

The MAC sublayer is responsible for scheduling resources, multiplexing channels, and error correction.

The main MAC Sublayer functions include:

- Mapping (logical channels and transport channels).
- Multiplexing and demultiplexing Media Access Control Software Data Units that belong to one or a different logical channel within the transport block are delivered on the transport channel to and from the physical layer.
- Reporting of Scheduling Information.
- Hybrid Automatic Repeat Request error correction.
- Employing dynamic scheduling
- Logical channel prioritization and handling.
- Padding.

4. C-RAN

The ever-increasing need for faster data transmission has introduced more challenging CAPEX and OPEX expenditure in the Radio access subsystem of mobile systems. Network operators must increase network bandwidth to increase data transfer rate and volume. An access network consists of base stations and base station controllers for transmitting signals and data. This method is the most common and traditional way to deploy the access network. These components require huge expenditures such as hardware, software, electricity, supporting equipment, and site costs.

Organizations require modern technologies such as MIMO and infrastructure antennas to increase network capacity.

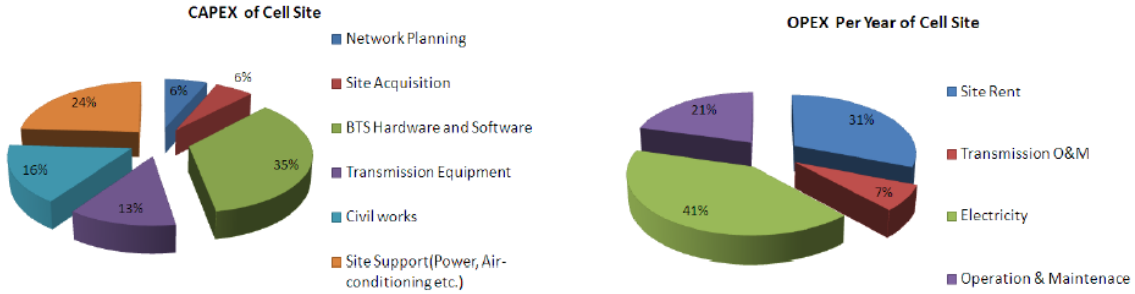


Fig. 37. CAPEX and OPEX Analysis of Cell Site [19]

The total cost of ownership (TCO) of a mobile network includes :

- Operating Expenditure (OPEX)
- Capital expenditure (CAPEX).

Approximately 60% of the total cost of ownership is operating expenditure, and 40% is capital expenditure. CAPEX consists of expenses related to cell site construction, including charges from network planning to site acquisition. Acquisition of base stations and additional equipment such as power, cooling, and deployment. It also includes construction costs like installation, procuring leased lines, software licenses, and radio frequency types of equipment. OPEX costs are related to network operation and maintenance, i.e., electricity, leased line, site rental, and upgrade. The more base stations are installed, the more CAPEX and OPEX will increase. Specifically, base stations increase capital costs as they are the most expensive wireless network infrastructure components. Cell sites increase operating costs because cell sites require significant amounts of power to operate. Therefore, new and novel architectures that optimize cost and power consumption are becoming essential in mobile networks. [19].

4.1.1 Challenges of RAN

- High power consumption due to a large number of base stations installed
- Rapid increasing capital and operational expenditure of RAN
- Mobile network load
- Low base station utilization rate
- Network capacity need

4.2 C-RAN Architecture

Under Wireless Network Cloud (WNC), IBM was the first to introduce the Cloud Radio Access Network (CRAN). Later, it was described in detail by the China Mobile Research Institute and others. It was primarily done to resolve the cost issue at hand in the traditional radio access networks. The CRAN architecture provides improved performance, easy scalability, flexibility, and lower mobile network costs.

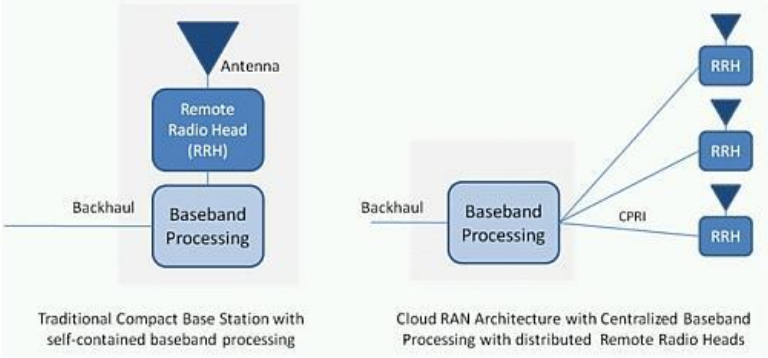


Fig. 38. Traditional base station model and C-RAN model [19]

In a classical RAN, all the eNodeB radio components are located at each cell site consisting of two elements :

4.2.1 Remote Radio Head (RRH)

- Consist of transceiver or antenna for the mobile device.
- Performs digital processing
- Power amplification and filtering

4.2.2 Baseband Unit (BBU)

- It is used to process call and forward traffic to the mobile switching center.
- Preforms signaling to RRHs.

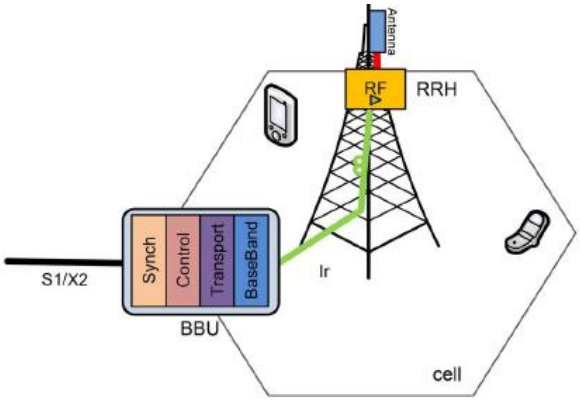


Fig. 39. Traditional Base Station with RRH [20]

These components are interconnected by a common public radio interface (CPRI). In addition, the X2 interface is defined between base stations. Finally, the S1 interface connects a base station with the mobile core network.

BBUs have been moved from individual cell types to a centralized pool of BBUs over high-bandwidth links in CRAN. Performing baseband processing requires general-purpose processors configured through a baseband unit virtualized cluster. [20].

A wide range of radio access networks for coverage and high capacity is provided by distributed RRHs. Due to the small size and ease of installing the RRH, the CAPEX and OPEX are much lower and can be used in relatively dense and large areas. On the other hand, all radio access technologies must connect to the baseband unit over a low-latency, efficient bandwidth optical transport network. High-performance generic processors are used to build these baseband processing pools. It aggregates the processing power and provides the signal processing capacity required for virtual base stations in the pool through real-time virtualization technology. Centralized BBU grouping dramatically reduces the number of BS rooms required, allowing large-scale coordinated radio transmission and reception with resource aggregation. [21].

Centralized topology simplifies network deployment and accelerates scalability. One baseband unit can serve many remote radio heads. The distance between a baseband unit and remote radio head can span up to 40 KM using fiber optic and microwave links.

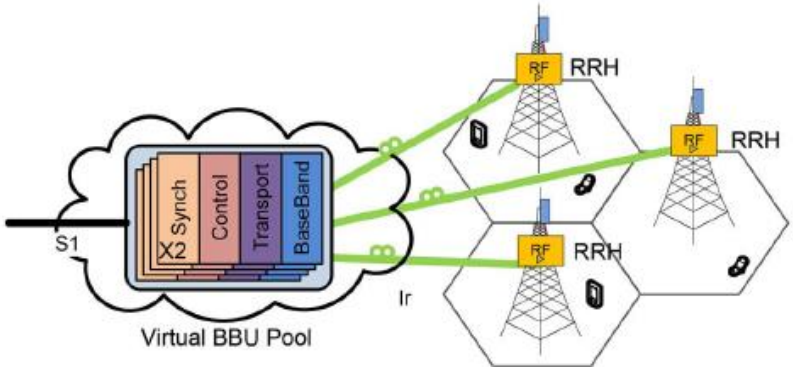


Fig. 40. C-RAN with RRHs [20]

CRAN can have multiple meanings and be interpreted as coordinated, centralized processing, cloud, clean, etc., and most widely used to describe architecture.

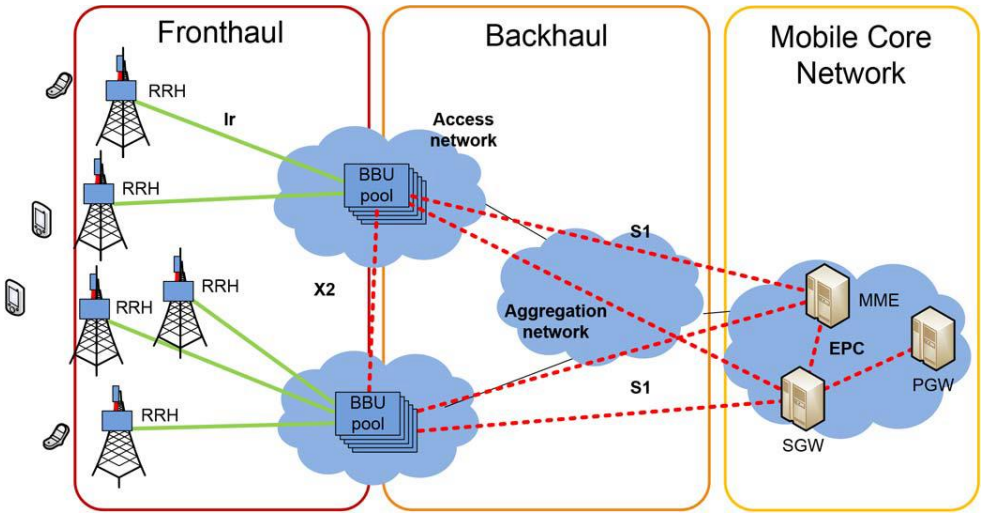


Fig. 41. C-RAN architecture [20]

The baseband unit and remote radio head are connected through a front link, and the Baseband unit and the core network are connected through a backhaul link. The fronthaul architecture uses the Ir interface to connect the RRH to the BBU. The backhaul part uses the S1 interface to connect the BBU pool to the mobile core network and the X2 interface to connect the BBU pool. The RRH at the remote site is deployed using an antenna. RRHs in the BBU pool are connected to high-performance processors through an optical transport link.

Architecture	Radio and baseband functionalities	Problem it addresses	Problems it causes
Traditional base station	Co-located in one unit	N/A	High power consumption Resources are underutilized Resources are underutilized
Base station with RRH	Spitted between RRH and BBU. RRH is placed together with an antenna at the remote site. BBU is located 20-40 km away. Generally deployed nowadays	Lower power consumption. More convenient placement of BBU C-RAN	Resources are underutilized
C-RAN	Spitted into RRH and BBU. RRH is placed together with an antenna at the remote site. BBUs from many sites are co-located in the pool within 20-40 km away. Possibly deployed in the future	Even lower power consumption. The lower number of BBUs needed - cost reduction	Considerable transport resources between RRH and BBU

Table 5. Comparison between a traditional base station, base station with RRH and CRAN [20]

4.2.3 Advantages of C-RAN

- Reduced costs and provided scalable resources.
- On-demand resource allocation
- Co-operative processing and resource sharing
- Deployment of advanced technologies
- Joint transmission and processing
- Adaptability to Non-uniform Traffic

5. Network Function Virtualization (NFV)

The base station controller controls the base transceiver station and is connected to the mobile switching center in previous generations of cellular networks, such as 2G and 3G. These network elements are now designed with specialized hardware, i.e., network nodes with dedicated hardware. In a non-virtualized environment, the implementation of network components such as hardware and software is vendor-specific.

So as a result, the required hardware was expensive to procure, install and maintain. Furthermore, if there was a fault with a network node, the hardware eventually needs to be replaced, e.g., the base station controller had to be replaced by the same vendor that initially built it. This trend led to the monopoly of some vendors, and in turn, the cost of equipment skyrocketed and made it difficult for new vendors to make an impact in the telecom industry.

Another issue with network capabilities, integration, and upgrades was that it was hard to incorporate new use cases of technology such as the Internet of Things with traditional approaches since it was difficult to change the hardware design.

The solution proposed to the problem is the approach of network function virtualization (NFV). Network functionality is implemented in the software, and that software runs on commercially available hardware (COTS) such as servers, storage devices, hard drives, network switches, etc., that are not specialized. They are readily available in the market.

5.1.1 Objectives [22]

- Compared to dedicated hardware implementations, capital efficiency has increased. It is accomplished by employing commercial off-the-shelf hardware, such as storage devices and servers to offer virtualization technology to network operations. These network functions are called virtualized network functions (VNFs). In addition, reducing the number of different hardware architectures and sharing them also contributes to this objective.
- More flexibility when assigning virtual functions to hardware.
- Provides scalabilities and decouples functionality.
- Deployment of software-based service provides innovative rapid service.
- Migrating workloads and shutting down unused hardware to reduce power consumption.
- A standardized interface between the infrastructure and VNFs with related management entities.

The hardware is executed with specialized running software through which network functions are implemented as software processes. In other words, it separates the software from the hardware. Separating hardware from software allows flexible network functions and will enable resources to be shared and redistributed to perform different functions from time to time.

5.1.2 Advantages

- Both hardware and software costs are reduced.
- New vendors to enter the market.
- Easy to upgrade the network by simply upgrading the software.

5.2 NFV Architecture

The main domains are :

5.2.1 Virtual Network Function:

Software implementation of network functions.

5.2.2 NFV Infrastructure:

It Comprises physical hardware resources, such as computers, storage, and networking devices. The resources are logically partitioned and abstracted through a virtualization layer.

5.2.3 NFV Management and Orchestration

It comprises physical resource management, virtual network functions, and virtual resource management.

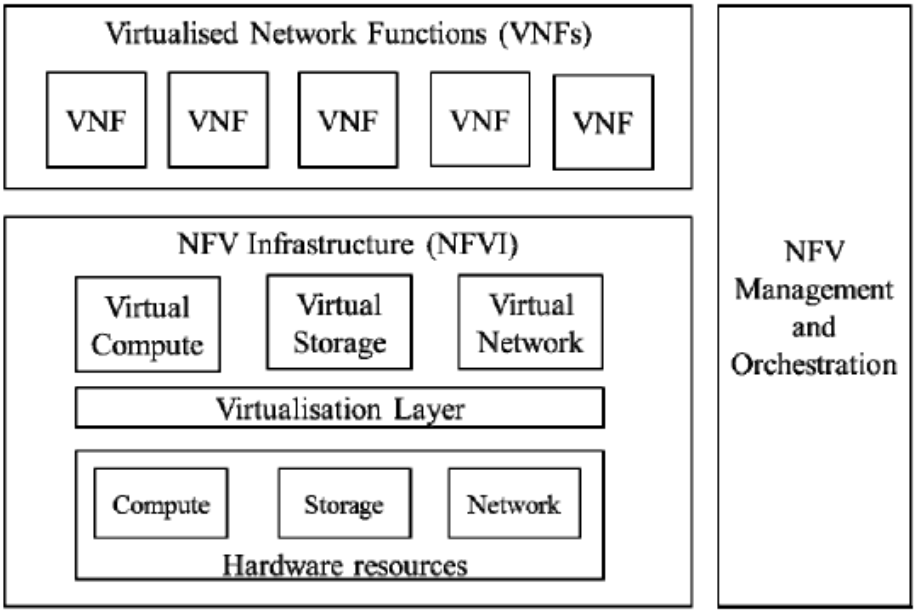


Fig. 42. NFV Architecture by ETSI [22]

As shown in the above diagram, the ETSI (European Telecommunication Standards Institute) presents an overview of the NFV architecture, with virtual network functionalities implemented in software at the top layer. They run of the virtualized infrastructure consisting of commercial off-the-shelf hardware resources, general processors, hard drives, and switches readily available in the market. The virtualization layer transforms these COTS into virtualized interfaces that serve as virtual computing, storage, and networking resources. These software processes that implement network capabilities are executing on top of these virtualized sources. All physical resources, virtual resources, their partitioning, and VNFs, all of this is managed by the network management and orchestration function. This function determines the number of resources allocated to the software process and the logical partitioning of these virtual resources.

5.3 NFV Reference Architecture

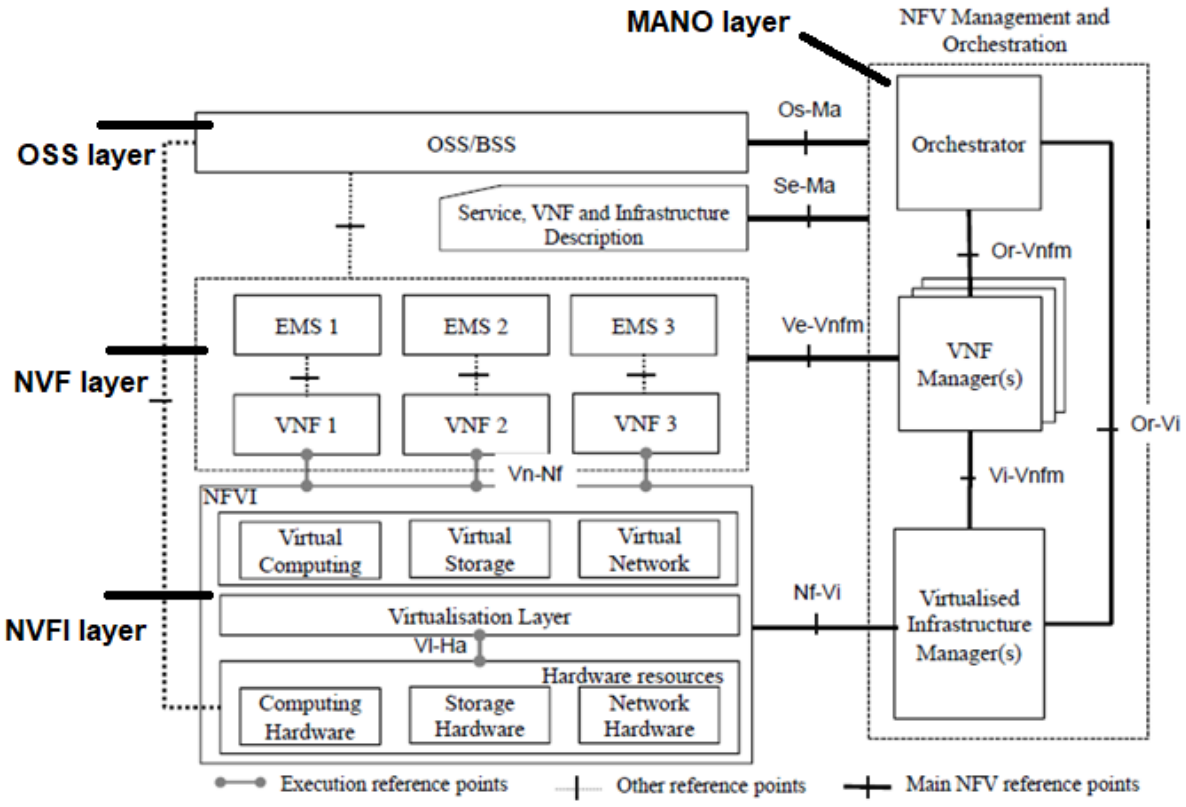


Fig. 43. NFV Reference Architecture by ETSI [22]

NFV architecture can be divided into the following four parts :

1. **Virtualization Network Function (VNF) Layer**
2. **NFV Infrastructure (NFVI) Layer**
3. **Operation Support Subsystem (OSS) Layer**
4. **Management and Orchestration (MANO) Layer**

5.3.1 Virtualization Network Function Layer (VNF)

It consists of two components :

5.3.1.1 Virtual Network Function (VNF)

Its purpose is to virtualize the underlying resources, and it is the fundamental building element of the NFV architecture. For example, a virtualized router or base station is referred to as a VNF router or a VNF base station. Therefore, when a function is virtualized, it is called a Virtual Network Function (VNF).

VNFs are distributed and run on a single or several virtual machines (VMs), each hosting a single component. Each has its VNF, and they function as a virtualized EPC together. Example: EPC virtualizes sub-functions like HHS, MME, and gateway.

5.3.1.2 Element Management System (EMS)

Responsible for VNF function management and can be deployed as one EMS for each VNF or as one EMS for multiple VNFs, including :

- Configuration
- Security
- Accounting
- Fault
- Performance

5.3.2 NFV Infrastructure Layer (NFVI)

It is a collection of hardware and software components that assemble an environment in virtualized infrastructure that manages, deploys, and executes virtual network functions. The network provides communication between the NFV infrastructure's various locations. The NFV infrastructure can physically span multiple locations [22].

The following are included in NFV Infrastructure :

- Hardware Resources
- Virtualization Layer
- Virtual Resources

The virtualization and hardware layers are considered a single unit from a VNF perspective.

5.3.2.1 Hardware Resource

Includes :

- Storage, computing, and network that provides processing.
- VNFs connectivity through Virtualization Layer (hypervisor)

Computing storage and hardware are assumed to be COTS and be pooled together. Network resources consist of switching functions such as routers and wired or wireless links.

5.3.2.2 Virtualization Layer

The virtualization layer decouples VNF software from hardware, providing an independent VNF lifecycle and abstracting hardware resources. In addition, it is responsible for the following [18]:

- Physical resources are separated logically and abstractly via a hardware abstraction layer.
- This software implements virtual network functionality and leverages the virtualization infrastructure.
- Provides virtualized resources for VNFs to execute

5.3.2.3 Virtual Resources

The hardware tier's computing, storage, and network resources abstraction are virtual resources.

Operation Support Subsystem (OSS) and Business Support System (BSS) Layer

OSS is responsible for network management, fault management, configuration management, and service management, whereas BSS is responsible for customer management, product management, order management, etc.

5.3.3 Management and Orchestration Layer (MANO)

It comprises three components [22] :

5.3.3.1 Virtualized Infrastructure Manager(s)

It includes the features used to control and manage the interaction of VNF with computing, storage, and network resources under its privileges and virtualization.

Virtualized infrastructure Manager performs the following :

- Computing, storage, software, and network resources specifically for the virtualized infrastructure.
- Manage infrastructure resources and allocations. For example, to increase the number of virtual machines, energy efficiency, etc.

- Virtual machines assign hypervisors with storage, computation resources, and network connections.
- Root cause and analysis of performance issues from an NFV infrastructure perspective.
- Collecting infrastructure fault information.
- Information collection for capacity optimization, planning, and monitoring.

5.3.3.2 VNF Manager

Responsible for lifecycle management of VNFs, including installation, updates, query, and termination of VNFs.

5.3.3.3 Orchestrator

Controls software resources and realizes network services. It is also in charge of the NFV infrastructure and orchestration.

5.3.4 Advantages

- 5G is a collection of network functions that are virtualized and can run on a cloud infrastructure.
- Cost-effective network based on COTS
- Ability to deploy new network functions quickly
- Shared usage of the resource
- Scale up or scale down resources as per requirement. Increases AMF output in minutes using MANO

6. Software Defined Networking (SDN)

There are two levels of abstraction in a distributed computer network :

1. The control plane is responsible for constructing, implementing, and maintaining the routing table.
2. The forwarding plane is also known as the data plane responsible for forwarding the data packets to their intended location.

In an era with massive implementation of computer networks consisting of a wide range of routers and switches, the complexity of managing and implementing functionalities by network operator on each device increases exponentially as different manufacturers and vendors provide the device with specific commands. In addition to maintaining configuration consistency, policy formation, fault tolerance, and load changes on individual routers become problematic because they have to be done manually on each device. Due to the size, heterogeneity, and complexity of distributed computer networks, traditional optimization, configuration, and troubleshooting approaches would not be efficient and insufficient in some cases. [23]

The research community and industries had noticed the problems mentioned above to develop a better technological solution. Some of them are Named Data Networking (NDN) and Software-Defined Networking (SDN).

Software-defined networking (SDN) is the most widely accepted model for changing the boundaries of distributed network infrastructure. The Open Networking Foundation (ONF) is a non-profit consortium dedicated to developing software-defined networking (SDN) and standardizing the OpenFlow protocol.

ONF defines SDN as the physical separation of the network forwarding plane from the control plane, where several devices are controlled by the control plane.

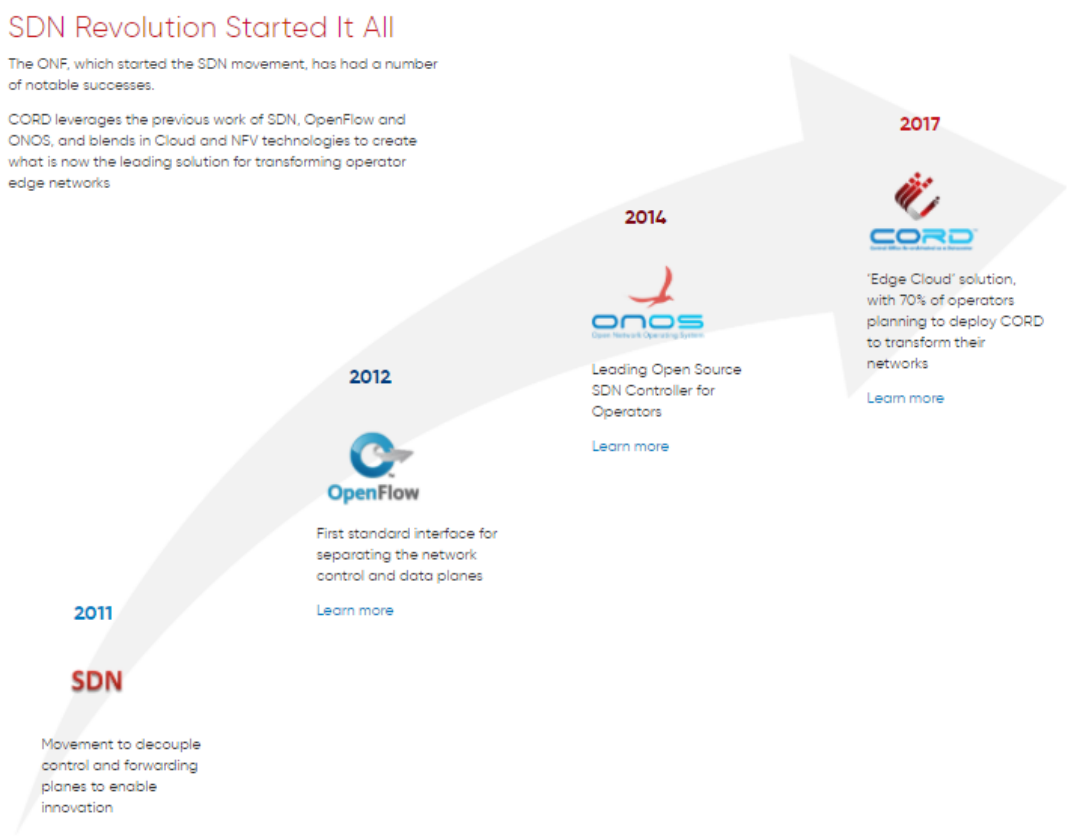


Fig. 44. Evolution of SDN and ONF [24]

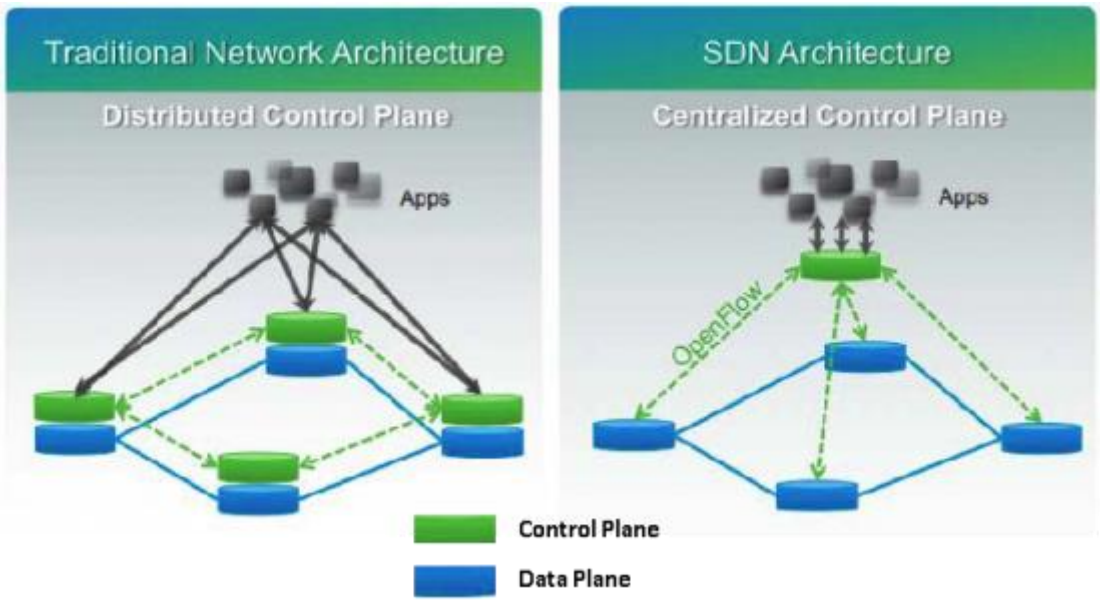


Fig. 45. Traditional Network vs. SDN architecture [25]

6.1 Principle Of SDN

The following are the principles of SDN as defined by ONF [26] :

- Controller and data planes decoupling -
  This approach requires a separate controller and data planes. First, however, it is essential to understand that control must be performed within the data plane system.

- Centralized logical control -
  Centralized controllers have a broader perspective of the resources under their control than local controllers and may make better deployment decisions. Both separation and centralization of control improve scalability. It will increase the globalization of network resources, but with minor detail.

- Expose abstract network states and resources to external applications.

6.2 SDN Reference Model

Terms and definitions [26].

6.2.1 Data plane

A collection of network elements, each containing a collection of resources for processing or forwarding traffic, comprises the data plane.

6.2.2 Controller plane

The controller plane consists of a set of SDN controllers, each controller exclusively controlling a collection of resources provided by one or more network elements in the data plane.

6.2.3 Application plane

The application layer consists of many applications having exclusive control over all the resources provided by one or more SDN controllers.

6.2.4 Management

The manager's most basic purpose is to establish reachability information, which allows lower- and higher-level plane entities to communicate with one another and transfer resources from the lower-level resource pool to specified higher-level client entities. A manager's functional interface exists for each SDN controller, network device, and application.

6.2.5 SDN controller

It's a software program solely responsible for data plane resources and abstraction. At least one client can get the abstract information model from the SDN controller.

6.2.6 SDN controller interfaces

- Application-controller plane interface (A-CPI) :
  It establishes a link between the application and the SDN controller, allowing it to receive services from the SDN controller.
- Data-controller plane interface (D-CPI) :
  Provides connection between the data planes and the controller, through which the SDN controller controls the data plane resources.
- Management interface :
  Through which resources and policies can be established and other traditional management functions.

6.2.7 Network element

Defined as a group of data plane resources managed as a single unit.

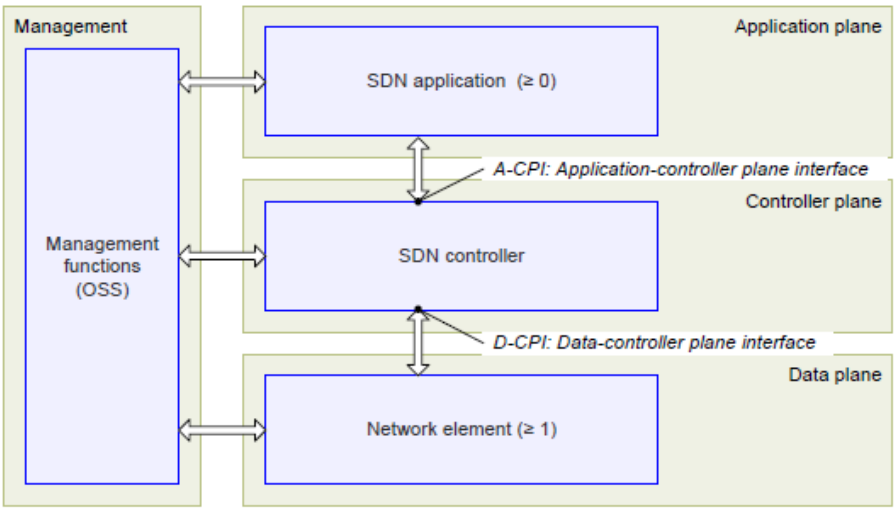


Fig. 46. SDN reference model [26]

At the bottom, we have the infrastructure layer or the data plane consisting of network elements like switches, servers, and edge routers for traffic forwarding based on rules provided by a controller and have minimum control and management functions. In between, we have the controller plane or layer where all the controller's functionalities reside and centralized acting as the brain. The SDN controller can have a different module to provide the network elements with logic to control, translate the application requirement from the application layer, and provide a programmable platform for the SDN applications. It also has plugins to interact with data and application plane via interfaces. The Northbound Application Controller Interface connects the application plane and the controller plane, and its purpose is to give network status or import forwarding rules. The South Data Controller Interface is the interface that connects to the data plane, and its purpose is to offer an access point. An east-to-west interface may also transmit information and coordinate decisions across multiple controllers in an extensive network [26]. Finally, at the top, we have the application layer where SDN applications are designed to fulfill user requirements and communicate their network requirements toward the SDN controller.

7. Cybersecurity In 5G

Recent advances in mobile communications have created security challenges that have significant privacy concerns for commercial and industrial 5G applications. Many of the security concerns associated with the introduction of 5G are addressed through security measures.

For a long time, business and academic researchers have been striving to improve 5G security. Since 2017, a 3GPP working group dedicated to service and system aspects has been researching and establishing security specifications for 5G systems [27].

The 5G network, according to 3GPP, is divided into two parts:

1) Standalone Network
2) Non-standalone Network

7.1 5G NSA

In this scenario, the evolved packet core is operational, which is the core of 4G networks, and prefers E-UTRAN, the access network of the 4G LTE. As a result, it is crucial to remember that threats and vulnerabilities in 4G LTE networks can equally affect 5G networks.

The following are the critical threats to the security of the 5G NSA network [27]:

7.1.1 Downgrade Attack

It forces the UE LTE connection to 2G or 3G, but the end-user can connect using more advanced technology. Ultimately, the attacker could perform a man-in-the-middle or eavesdropping attack to collect information. For example, a customer can identify whether they have an LTE connection at a location, and it suddenly drops to "E," "G," or another symbol by looking at the indication on their devices.

7.1.2 Data modification Attack

Secure methods of intercepting traffic do not protect the integrity of UMTS and LTE communications. It can result in active men-in-the-middle such as data injection or modification. Mobile devices and base stations can prevent man-in-the-middle type attacks by authenticating and verifying each other. The 5G protocols such as AKA and EAP-AKA initiate the authentication process on 5G networks and are new solutions for recording connection requests

7.1.3 IMSI Tracking

The International Mobile Subscriber Identity (IMSI) is transmitted wirelessly, unencrypted, allowing an attacker to find the SIM card used by the connected user when IMSI (International Mobile Subscriber Identity) requests are created. Additionally, base station spoofing is a fake base station that can unknowingly track and collect personal data.

7.1.4 LTE Roaming

Vulnerabilities in 2G, 3G, and 4G users may be exposed to attacks such as eavesdropping on calls, reading and forwarding data, and tracking due to the usage of obsolete signaling protocols such as SS7 for PSTN networks and diameter for authentication and authorization.

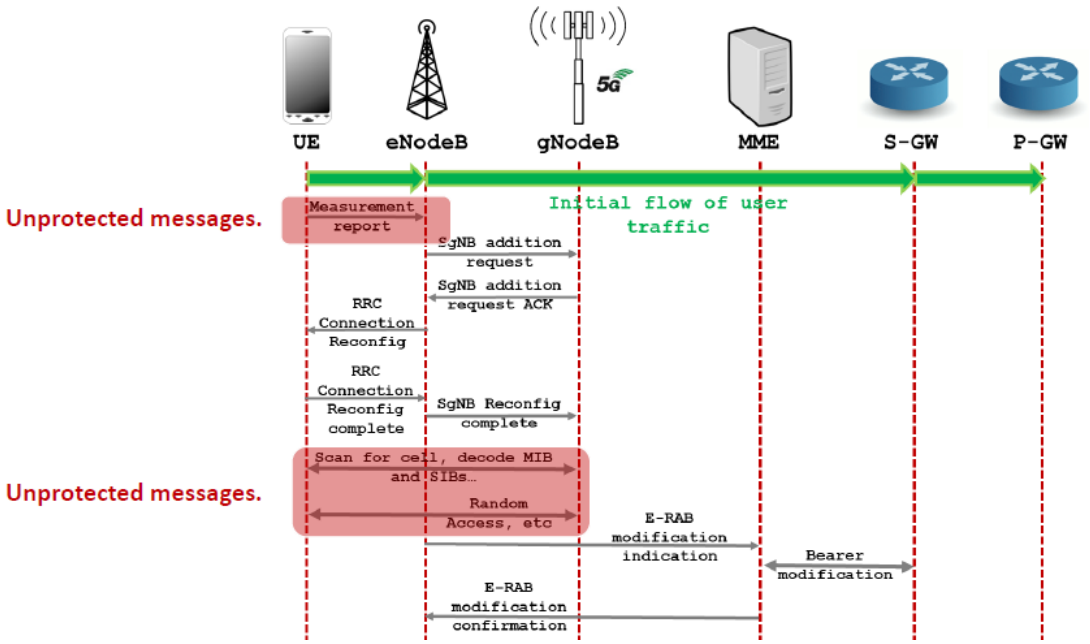


Fig. 47. 5G NSA Attach Procedure [28]

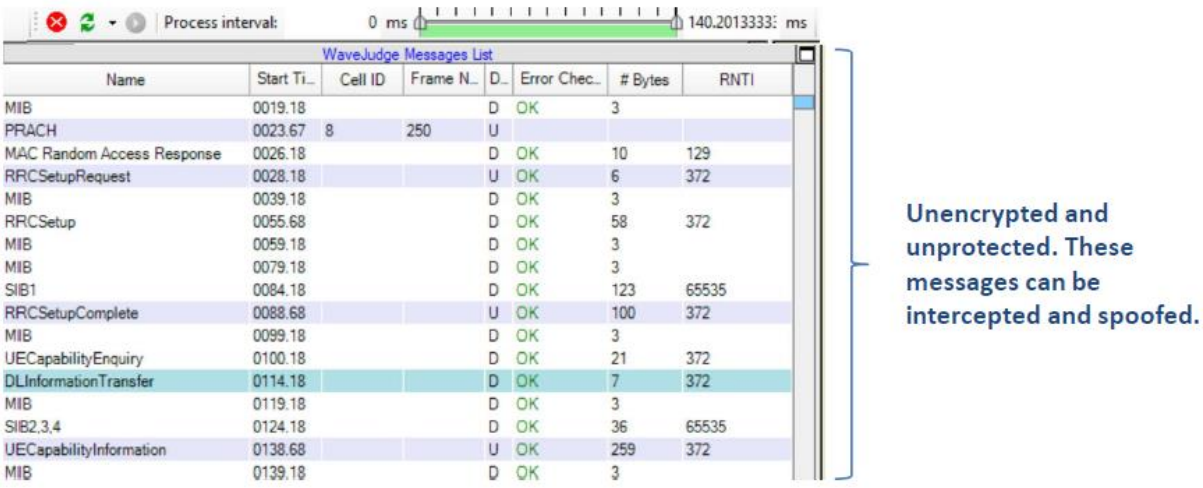


Fig. 48. Message list [28]

7.2 5G SA

The service-based architecture of 5G SA, which uses technologies like NFV and SDN to boost security solutions, is a significant feature that distinguishes it from 5G NSA. The service-based architecture runs on general-purpose hardware with software applications built on it. The open-source nature creates vulnerabilities. To address security concerns with 4G LTE, the new ITU standards are included in 5G stand-alone services, protocols, and implementations. The information transfer is software-based in 5G, which is sensitive and confidential, poses a risk to the network applications if not targeted with proper authentication and encryption.

7.3 Threats/ Vulnerabilities In 5G SA

7.3.1 SUPI/SUCI Privacy

Subscription Permanent Identifier (SUPI) is encrypted to Subscription Concealed Identifier (SUCI) and transmitted over the network in 5G, preventing attackers from targeting and tracking subscribers' privacy.

SUPI can be targeted and monitored over the air under specific circumstances [29] :

- An unauthenticated device is trying to make an emergency call.
- UE with outdated SIM not provisioned with 5G public key.
- Customers bring their device with an un-updated SIM card to the operator's network.

7.3.2 Man-in-the-middle (MitM)

In 5G, there is a significant concern related to the unprotected user plane integrity protection that an attacker can exploit with Man-in-the-middle during its registration on the network. IP is enabled on control plane messages, but the data plane and control plane are separated, leaving the data plane vulnerable.

The three classes of attacks that can occur are [27] :

- Identification attacks, detection of devices on the network, knowledge of their properties, and applications.
- Reduce attacks and data rates with capturing device capabilities.
- Battery drain attacks.

7.3.3 Roaming

Roaming from 5G SA to NSA is a network security issue. Even now, the NSA, like 4G EPC, uses SS7 and protocols that are vulnerable to attacks such as text message decryption, location monitoring, and eavesdropping. Therefore, protocols such as HTTP/2 and JavaScript Object Notation (JSON) should be used to mitigate these vulnerabilities with roaming flows.

7.4 Threats Related To gNB

7.4.1 Spoofing

Rouge gNBs might have a significant impact depending on the network design of the operator, e.g., the interface between gNB and AMF is protected and secured by IPSec. However, installing a fake gNB is easy and cheap using a software-defined radio-based solution. Suppose the gNB cannot access and penetrate the core network of the compromised network operator. In that case, the effect is limited to misconfiguration of the UE attempting to register or re-register with the core network via this gNB. In these rare cases at the protocol level, previously reported DoS attacks on UEs in 4G networks are common. [17]

The fake gNB can still access the KgNB associated with the UE and integrate itself into the core network, and the UE will not be able to identify the gNB as a fake. User plane encryption is done directly between the gNB and the UE, so knowing the KUPenc key is also required. Therefore, If end-to-end encryption is not provided at the application layer between the UE and the final data sink, the user plane data is mapped directly to the gNB level. However, knowledge of KgNB is still insufficient and does not imply knowledge of KNASint or KNASenc. Therefore, encryption should be enabled by AMF at the NAS level so that gNB cannot generate or intercept NAS messages that request or include User PII such as IMEI or PEI [17]. Remember that identification request messages at the NAS level (response may contain PEI) are sent only after successful authentication if the security context has been established at the AMF level.

7.4.2 Tampering

A software update method is to be used to update a gNB. For example, if the gNB firmware contains a backdoor and an attacker intentionally or unintentionally inserts the backdoor. At the same time, the debug feature is enabled, the modified node can violate the security features of security like IPSec and can lead to leakage of user secrets. IPsec implementation must meet specific requirements compared to long-term 5G device keys. However, the standard does not specify the minimum level of security. After removing the IPSec protection layer, it is also possible to use the decrypted data to access the storage directly.

7.4.3 Jamming by rogue gNB

A rogue gNB with a higher link budget will try to connect to UE, causing the legitimate gNB's connection to be disrupted. As a result, registration may fail. A rogue gNB may indicate that the user is unauthorized while transmitting a denial of connection with a reject cause, possibly resulting in roaming not permitted update status. Generally, the UE will not register until the device is powered off or the SIM card is removed and reinserted. The communication interface can be disconnected entirely for machine-to-machine communication for both fixed and mobile use if the device is altogether standard compliant. The registration with the new gNB is likely for mobile devices if the link budget with the existing serving cell has changed recently and for stationary if the link budget has not changed. [17]

8. Security Challenges In 5G RAN Technologies

8.1 Security Challenges For C-RAN And Related Technologies.

In today's telecommunication advancement, commercial networks and industries are booming to the new access technologies associated with 5G. However, it has generated substantial risks and challenges in the case against cyber-attacks.

As we know, absolute security protection from vulnerabilities is a myth. Still, in the case of the telecommunications sector, they may be avoided from the moment of initial entrance into the network before it gains control of the hardware. It is tough to identify attacks, as attackers use different patterns and frequently change their attack modes and methods. A single information breach can have a massive impact on the overall network and can compromise the sensitive information of customers and service providers.

C-RAN has led the telecommunication industry to satisfy the needs and demands of end-users for higher mobility and seamless connectivity. CRAN features centralized processing and control, provides collaborative access to wireless networks, and integrates with cloud-based systems. [30]

In 2015, [31] provided a logical abstraction of the C-RAN architecture, including physical plane, control plane, and service plane. It focuses on service-oriented cloud architecture, commerce, and the scheduling and management of personal resources [31].

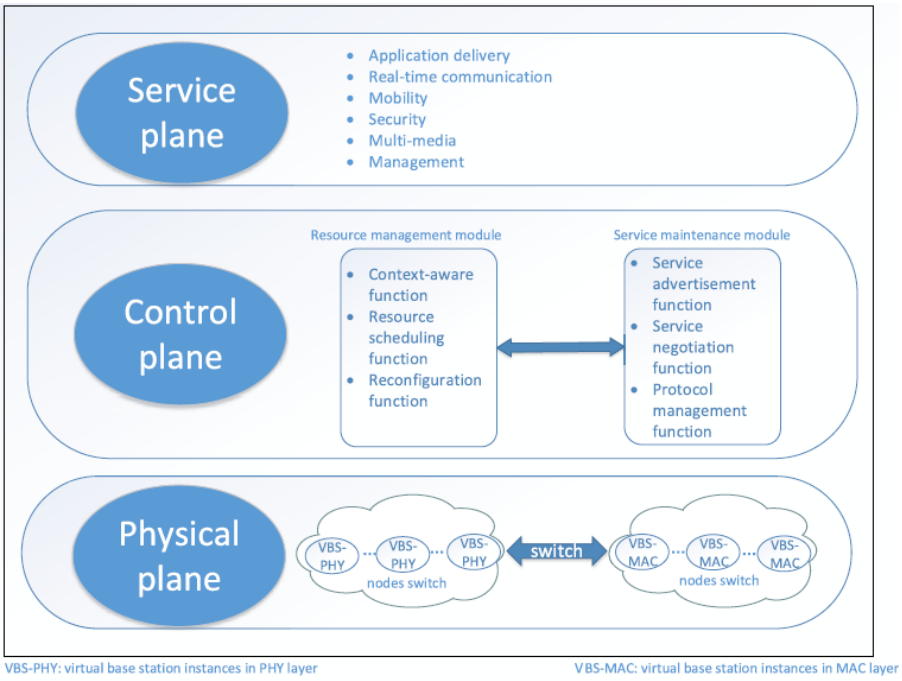


Fig. 49. C-RAN logical architecture [31]

This section focuses on security and the physical layer threats responsible for sending and processing signals from various nodes, communicating, and performing resource allocation. Physical plane security is the basis for ensuring a secure and reliable CRAN system. Consequently, this plane has been a focus of security concern [31].

8.2 Attacks And Threats

8.2.1 Eavesdropping Attack

The most popular method of stealing essential data and encrypted signals from wireless radio networks is eavesdropping. Massive MIMO is one of the key techniques of the C-RAN physical layer and has drawn attention for attackers on BS.

8.2.2 Jamming Attack

Jamming is a DoS attack that disrupts data communication and targets the physical layer. It can block user access, occupy most of the available spectrum, and jam the network.

8.2.3 Impersonation Attack

This attack can provide false information to nodes and refuse to provide service.

8.3 The Solution To Attacks And Threats

8.3.1 Eavesdropping Attack

Massive Multiple-Input Multiple-Output techniques can introduce two methods for detecting eavesdropping and active attacks to prevent BS from eavesdropping. The first method allows an authorized user to generate additional random PSK sequences, allowing base stations to intercept received sequences to detect attacks effectively. The second method eliminates the need for a different random sequence. In addition, beamformers are adapted to detect sniffing attacks.

8.3.2 Jamming Attack

SDN with centralized and separated planes can handle DoS attacks with application programming interface security.

8.3.3 Impersonation Attack

Allocation a default threshold for each node and obtain the suspicious level of a node by analyzing node reports. When the level of suspicion reaches a certain threshold, the node is considered a malicious identity node, and the information is excluded. This technique repeats this process for the remaining nodes until no malicious node is detected.

8.4 Security Requirements For C-RAN [31]

8.4.1 Access Control to Resources (Ac)

Unauthorized access to resources or services should be prevented at all times and in all places by the system.

8.4.2 Robustness (Rb)

Ensure the robustness required by users for cognitive radio channels to meet the QoS of communication.

8.4.3 Confidentiality, Integrity, and Availability (C/I/A)

The completeness of a system, its components, and any data or information transmitted to it are all referred to as integrity. All data, including user information and spectrum resources, must be treated with confidentiality and only accessible with authorization.

8.4.4 Authentication (Au)

The authentication mechanism permits the CRAN system to validate authorization and detect fake nodes and malicious users.

8.4.5 Privacy (Pr)

Most communication services are intended to gather data and personal information from the end-user. As a result, end-users may disclose privacy-sensitive information.

8.4.6 Trustworthiness (Tr)

Trust management mechanisms are becoming essential for implementing reliable cooperation between operators.

8.4.7 Compliant Towards Local Regulation Standard (CLRS)

The system must be developed to satisfy the local operator's regulatory standards, necessary for establishing a communication system.

8.4.8 Non-Repudiation (Nr)

This kind of action cannot be denied to verify any user's actions.

Attacks/Threats	C-RAN Security Requirements							
	AC	Rb	C\I\A	Au	Tr	CLRS	Pr	NR
Eavesdropping attack	N	Y	Y	Y	N	Y	N	N
CR node impersonation attack	Y	N	Y	Y	N	Y	N	Y
Primary user emulation attack	Y	N	Y	Y	N	Y	Y	N
	Y	N	N	Y	N	Y	N	N
	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	N	Y	N	N
	Y	N	Y	Y	N	Y	N	N
Wireless channels threats	N	N	Y	Y	N	Y	N	N
	N	N	Y	Y	N	Y	N	N
	N	Y	Y	Y	N	Y	N	N
	N	Y	Y	N	N	Y	N	N

Table 6. C-RAN security threats and requirements (YES/NO) [31]

8.5 C-RAN Technologies And Security Issues.

8.5.1 Massive MIMO

As we know, Massive MIMO combines many antennas to improve both the coverage and increase the data throughput. A large number of antennas at BSs can significantly improve the throughput performance and shift most of the signal processing and computation from user terminals to BSs. In addition, massive MIMO can improve communication security. However, eavesdroppers can utilize massive MIMO to attack legitimate communications [32]. If the eavesdropper's antenna array is much more powerful than the base station, it can decode the symbol. However, with the number of antennas at the BS sufficiently high, the massive MIMO eavesdropper fails to decode most of the original symbols. At the same time, legitimate users can recover the original symbols with only a limited number of antennas. [33]

8.5.2 Network Slicing

One of the main factors leading to network slicing threats is improper isolation between network slices and incorrect isolation between components within the same slice. For example, if one of the devices is infected with malware using a device vulnerability, the threat can be migrated between slices. As a result, other slices will also be impacted [32].

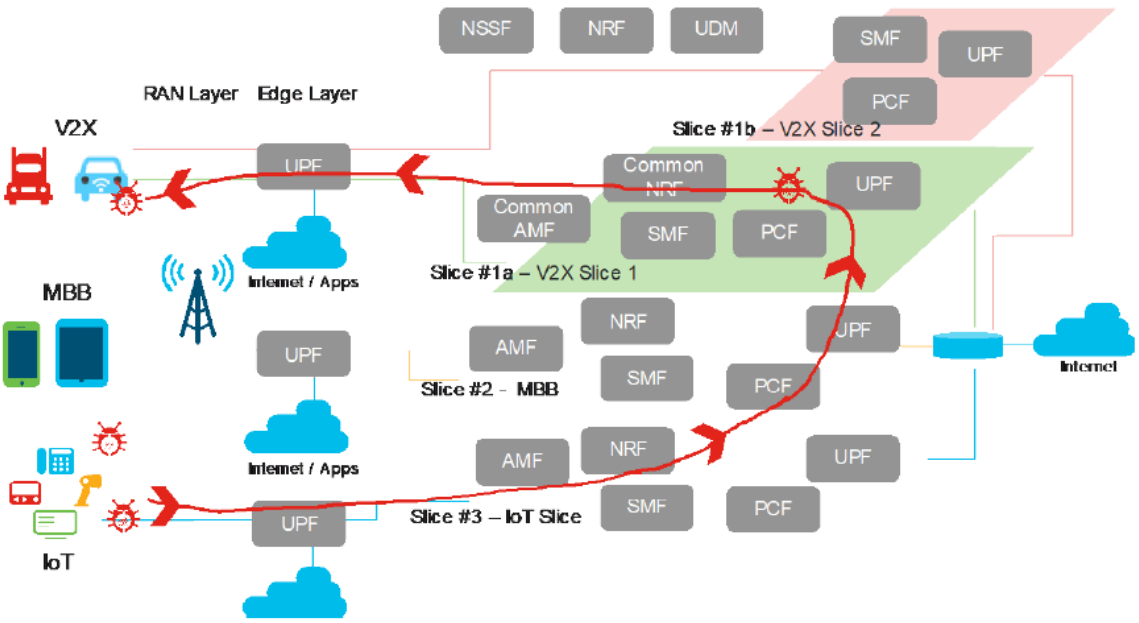


Fig. 50. Network Slice threat [32]

The attack can be extended and factorized by allowing the malware to exhaust slice resources, resulting in a DoS (Denial of Service) for the actual subscriber. An attacker can potentially exhaust resources across several slices, resulting in a denial of service or a significant decline of service in other slices. As a result, the network services provided suffer severely in quality. An attacker could also find a vulnerability in the compromised device or the endpoint terminal of other slices if the slice and its components are not configured and isolated separately

A high level of security that allows multiple logical networks to run as essentially independent business activities on shared physical infrastructure is required for a network slicing architecture. In particular, isolation between different layers of slices and their components can prevent vulnerabilities from reaching out to other slices and their components in the event of a malicious attack.

Two types of isolation can be differentiated in in-network slicing:

1. Resource Isolation
2. Security Isolation

The concept of resource isolation states that computing, storage, and network resources assigned to one slice of the network cannot be hacked or leaked into another slice. It ensures that when a slice is the target of a denial of service attack, the required slice resources are still available even when other slices try to scale and acquire additional resources [32]. In addition, isolation should be provided so that information in one part cannot be accessed or modified by other parts sharing the same infrastructure known as security isolation. Techniques such as software-defined partitioning, access control lists, and tagging can be employed to achieve slice isolation.

Isolation is the main feature provided by virtualization layers such as hypervisors. As a result, in the NFV environment, robust NFV security, comprising a reliable implementation of the virtualization layer and the entire cloud stack, can provide both forms of isolation. Even in distributed cloud deployments where SDN allows for highly dynamic and shared common transport infrastructure and flexible control over different virtual transport networks, a dedicated virtual transport network is still required to provide isolation. A transport network is still required for each slice to be generated for transmission between different hardware platforms. Firewalls are also necessary for security features like intrusion detection and anomaly detection, cryptographic encryption, and traffic monitoring, among others.

9. SDN Security and Issues

In Software-Defined Networking, the network control plane and data plane are separated, and a single control plane can handle several devices. There are, however, some security vulnerabilities that have yet to be addressed.

The SDN controller handles modifications and updates to flow rules in the data forwarding components. This control information traffic is easily identifiable, making it a visible network entity and a potential target for DoS attacks [34]. Additionally, it makes it easy for an attacker to obtain valuable confidential data with a compromised SDN controller quickly.

Similarly, in a congestion or saturation attack, centralized network management can also become a controller bottleneck for the entire network. Permitting programmability ensures that most network services can be implemented as SDN applications. If malicious applications are granted access or crucial application programming interfaces (APIs) are exposed to unintended software, havoc can be caused in the network. [35]

Data forwarding elements must store traffic flow requests until the controller updates the flow forwarding rules in the current SDN architecture (OpenFlow). As a result, data plane elements may be vulnerable to saturation attacks since forwarding elements, such as switches, have limited buffering capabilities.

In addition, since the architecture is controller-dependent, both the control and data plane channels must be resilient to security attacks, in contrast to the usage of current security protocol options and long recovery periods in large networks. Multiple controllers or redundancy can help with controller availability while improving security [34]. However, misconfiguration of forwarding elements or conflicts caused by several controllers, on the other hand, will make network-wide security enforcement difficult.

Targeted level	Malicious behavior	Caused by	Possible countermeansures
Forwarding plane	Switch DoS	Limited forwarding table storage capacity Enormous number of flows Limited switchs buffering capacity	Proactive rule caching Rule aggregation Increasing switchs buffering capacity Decreasing switch-controller communication delay
	Packet encryption and tunnel bypassing	Invisible header fields	Packet type classification based on traffic analysis
Control plane	DDoS attack	Centralization Limited forwarding table storage capacity Enormous number of flows	Controller replication Dynamic master controller assignment Efficient controller placement
	Compromised controller attacks	Centralization	Controller replication with diversity Efficient controller assignments
Forwarding-control Link	MITM attacks	Communication message sent in clear Lack of authentication	Encryption Use of digital signatures
	Replay attacks	Communication message sent in clear Lack of time stamping	Encryption Time stamp inclusion in encrypted messages

Table 7. SDN security issues with possible countermeasure [33]

9.1 Solutions For SDN Threats

SDN facilitates the identification of threats through a cycle of gathering intelligence from the networks as the system is centralized. As a result, the SDN may monitor security in both a reactive and proactive manner. It can also analyze traffic and respond to systems to assist network diagnostics [30].

The SDN architecture allows a proactive and highly responsive security monitoring system that analyses traffic and responds to help with network investigation, security service deployment, and policy modification. Packet resolution and flow are two of the significant characteristics of SDN since they provide transparency into a packet's origin or source, path, and even content. As a result, security applications can collect packet patterns through the control plane from any network perimeter to verify their content regardless of the network's ingress or egress ports, unlike traditional networks where security measures are usually located at entry points. This SDN feature provides a framework for network-wide security policies, early detection of threats anywhere on the network, and quick response by changing flow tables in real time to route threats to intrusion detection systems or firewalls. Security through SDN may be referred to as software-defined security because most security functions will be implemented in software at the application layer [34].

10. NFV Security and Issues

There are fundamental security challenges such as confidentiality, integrity, authenticity, and non-repudiation posed in NFV and are essential for future communication networks. One of the primary persistent challenges for using NFV in cellular networks is that the dynamic nature of virtual network functions (VNFs) leads to misconfiguration and thus to security breaches [35].

VNFs are also vulnerable to conventional cyber-attacks, including spoofing, sniffer, and denial-of-service attacks. NFVs are also vulnerable to a special set of virtualization threats, such as side-channel attacks, flood attacks, hypervisor hacks, malware injections, and migration-related attacks. Also, virtual machine migration, as well as cloud-specific attacks. Because remote access to the system is prohibited, private NFV implementations are only subject to attacks by malicious insiders, such as malicious administrators. A rogue user or a compromised VNF environment can introduce malware or alter network traffic to interrupt operations due to the apparent infrastructure accessibility.

10.1 Security Solution

The security of VNFs through a security orchestrator in correspondence with the European Telecommunications Standards Institute network virtualization function architecture is presented in the figure below. The proposed architecture ensures virtual functions in multi-tenant environments and secures the physical entities of the telecommunications network [34].

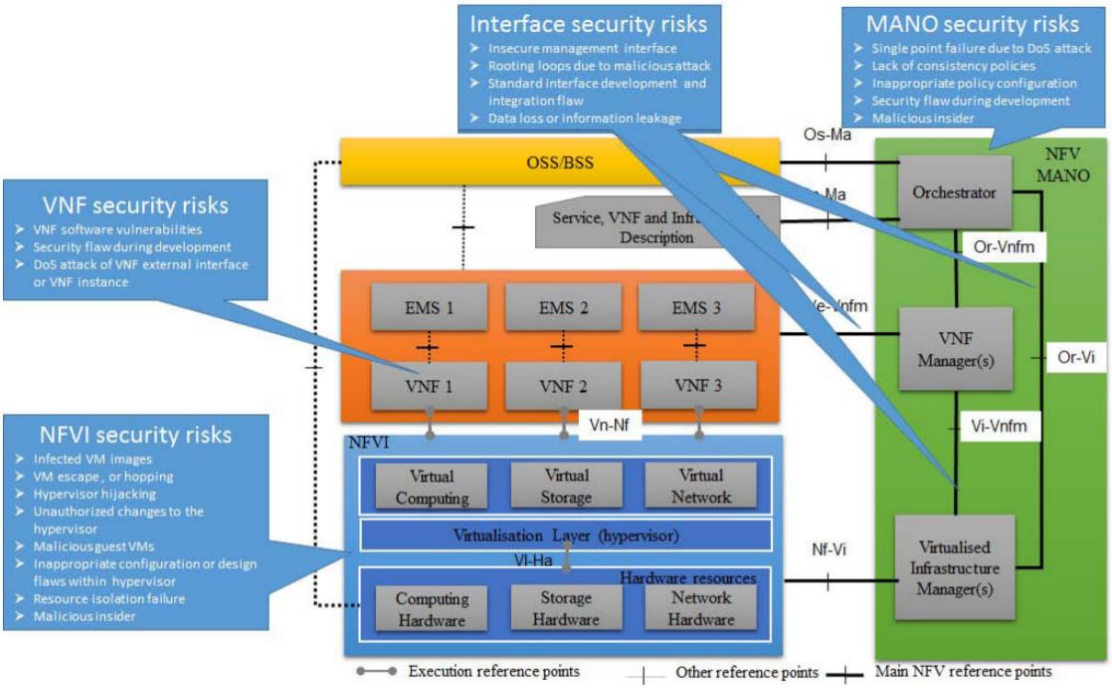


Fig. 51. Security Orchestrator NFV and risks [36]

Complex security solutions such as firewalls and IDSs can be used in NFV systems to prevent external attacks. In addition, identity and access control mechanisms such as role-based access control can be used to mitigate the effects of insider attacks. Infrastructure-level attacks can be prevented by continuously monitoring each user's resource consumption and blocking malicious requests based on a blacklist of IP addresses. To ensure the security of VNF, solutions based on encryption technologies such as message stream encryption can be used. To increase trust between different entities, a chain of trust relationships must be created and maintained in the NFV environment throughout its life cycle.

Additionally, VNF providers can use accountability and trust management to ensure that their software works without modifying the infrastructure provider's network. To transfer sensitive information to external networks, NFV provides secure outsourcing, which is another viable solution that will protect sensitive data and validate its integrity. In addition, migration mechanisms that maintain security establish secure interfaces with authorized source and destination parties, and malicious activity detection and reporting during migration are necessary to enable secure VM migration [34]. Also, using a hypervisor provides hardware protection for personal information and detects software corruption in a virtualized environment [37].

Firewalls and intrusion detection systems can also effectively filter traffic passing through an entry path, affecting the internal system switching table [30].

Security Technology	Primary Focus	Target Technology		
		SDN	NFV	Cloud
DoS, DDoS detection	Security of centralized control points	✓	✓	
Configuration verification	Flow rules verification in SDN switches	✓		
Access control	Control access to SDN and core network elements	✓	✓	✓
Traffic isolation	Ensures isolation for VNFs and virtual slices		✓	
Link security	Provide security to control channels	✓		
Verification for Integrity	Data and storage Security in clouds			✓
Service-based access Control	Service-based access control security for clouds			✓

Table 8. Security technologies and solutions [30]

11. Zero Trust

In 2010, John Kindervag coined the Zero Trust model; the traditional cybersecurity model is built on the concept of 'trust but verify. In many cases network operator grants certain devices, applications, and users as trusted and gives them unnecessary privileges to access other network parts. The fact that verifying once is not enough. It should be done on an ongoing basis for the so-called trusted devices, applications, and users claim to be. For this reason, Kindervag's preferred model is to 'verify and never trust.

With this strategy in mind, Kindervag's Zero Trust model has three key features [38] :

1. Secure access to all resources, regardless of location.
2. Access to everything on the network is based on a strictly enforced "need to know".
3. Monitoring detailed logs and inspecting all types of traffic.

This feature in practice means that :

Verify: authenticate and verify on an ongoing basis.

Give minimal/strictly limited access :

- Segment the network to create small control zones with gateways controlling and monitoring access between different components.
- Control access to the application, data, and resources such as personally identifiable information, sensitive intellectual property is stored and processed.
- Grant least privileged access based on requirements or roles.

Assume Breach: Plan as if the attacker is outside and inside the network

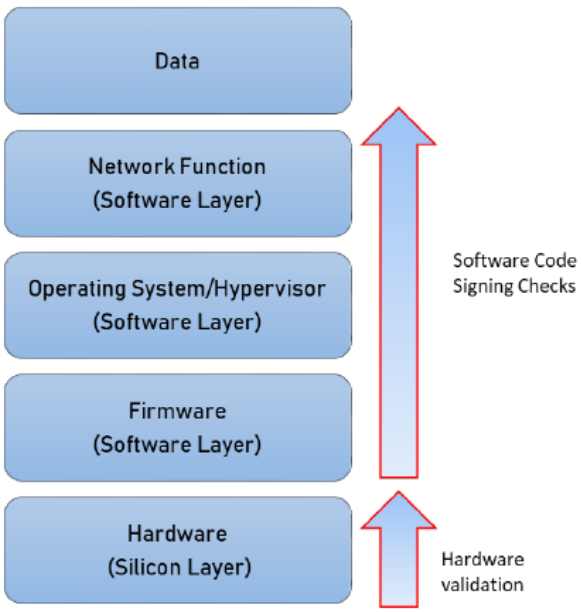


Fig. 52. Zero-Trust Validation checks at hardware and software layers [29]**.**

President Biden signed Executive Order 14028, "Improving the Nation's Cybersecurity" to Introduce the Zero Trust Model on May 12 to support our cybersecurity and protect the critical infrastructure and networks of the federal government that underlies our economy and lifestyle. Executive Order (EO) 14028, "Improving the Nation's Cybersecurity" introduces zero trust cybersecurity principles to authorities and adapt network architectures accordingly. To facilitate this effort, the Cybersecurity and Infrastructure Agency (CISA) has developed a zero-trust maturity model to help agencies implement a trustless (Zero Trust) architecture. This model complements the Office of Management and Budget's Zero Trust strategy. This strategy aims to provide the organization with action plans and resources to achieve an optimal zero-trust environment. [39]

An operative definition of zero trust and zero trust architecture is as follows [40]:

Zero Trust presents a collection of developed concepts and ideas to minimize the uncertainty associated with making accurate access decisions with minimal privileges on a per-request basis in information systems and services when faced with a network that is considered compromised.

Zero Trust Architecture utilizes the concept of zero trust and encompasses component relationships, workflow planning, and access policies in an enterprise cyber security plan.

As a result of the Zero Trust architectural design, Zero Trust is both physical and virtual network infrastructure and an operational guideline for businesses.

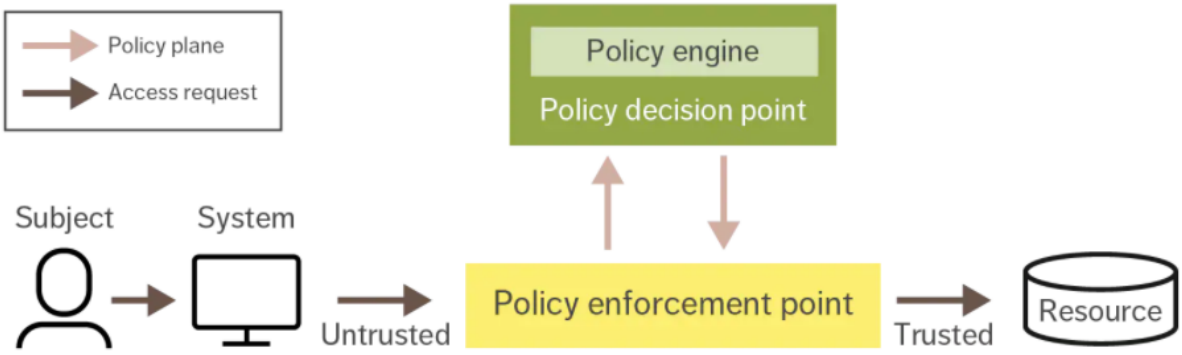


Fig. 53. Zero Trust Access [41]

The figure shows the logical components of the policy framework. Policy decision point (PDP) and policy enforcement point (PEP) are the essential logical entities. To gain access to a specific resource, the subject requests authorization at the policy decision time and provides information necessary for authorization and authentication. [41]

If adequately designed and implemented, zero Trust can help network administrators protect key systems and data from external and internal threat actors. For example, all attempts to access data are challenged and continuously validated in such an architecture, regardless of whether attempts to access the system and data are internal or outside the organization. [38]

Zero trust security is carried out from untrusted domains, such as user devices, the internet, the supply chain, service providers, and partners, to and from reliable and trustworthy domains, such as the operating company's network. Therefore, operators should be aware that Zero Trust can be seen as a security countermeasure for 5G.

12. Security As A Service (SECaaS)

5G offers a wide variety of service types associated with technical requirements. It also allows for various alternative use cases, each with security concerns. It thus requires the use of tools like NFV and SDN to provide differentiated services to customers. The operator's Security (SECaaS - Security-as-a-Service) offering is based on the network slicing, providing "security" for the use case and service exposure, providing service (as-a-Service). [29]

Service exposure covers aspects such as :

- Additional services such as security, geofencing, etc., for differentiation.
- Integration of slice properties such as latency, bandwidth, proximity, QoS, assessment, etc., is standardized.
- Insights and monitoring via instrumentation.

Network slicing is one of the enabling technologies that support 5G services. It allows every vertical service to have its network slice that provides the resources required for that vertical service. Network slicing is primarily based on SDN, NFV, and cloud computing. Creating different VNFs in different slices would create more vulnerability in other VNFs than in the static network. [42]

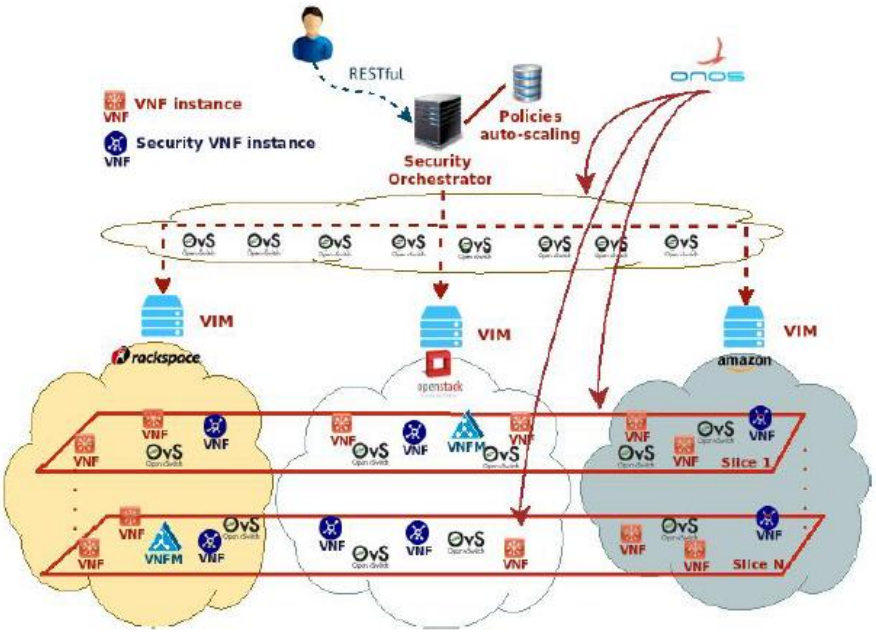


Fig. 54. Security as a Service architecture consideration [42]

Fig. shows an overview of the proposed architecture that enables SECaaS on cross-domain platforms. Various security VNFs, such as intrusion detection and prevention systems and deep packet inspection, will be deployed and managed using this architecture. The proposed architectural framework aims to dynamically deploy secure VNF instances maintain elasticity, monitor performance, and implement predictive auto-scaling based on pre-defined policies and metrics. [42]

12.1 Security As A Service With SDN

An intrusion detection system generates an alert and transmits it to security when it finds a malicious stream. Based on the number of alarms received and their severity, the orchestrator will direct the controller to temporarily or permanently terminate the malicious stream or avoid network overload while maintaining a particular level of service. The bandwidth is restricted. [42]

Algorithm 1 Attack-Response Algorithm
Require:
L: Level of the received alert.
T: Type of the received alert.
F: Flow that triggered the alert.

```
1: NumAlerts[L][T]   NumAlerts[L][T] + 1;
2: if NumAlerts[L][T]> = trigThreshold[L][T] then
3: triggResponse[F]   triggResponse[F] + 1 ;
4: end if
```

Fig. 55. Attack response algorithm for SDN [42]

12.2 Security As A Service With VNF

At each slice corresponding to the VNF manager, an auto-scaling algorithm is run to scale in and out each secure VNF instance based on established policies, as well as the VNF's performance and features. In addition, the auto-scaling solution must take into account the virtual machine start-up time as well as the number of virtual machines required and the data center load [42]

Algorithm 2 Scale-Out Algorithm
Require:
V ID: ID of the monitored VNF.
V|T: The type of the VNF.
FL: Flavor of the VNF.
SID: ID of the slice the VNF is assigned to.
CP: Type of the cloud platform on which the VNF is deployed.

1: **if** prediction(t0 + startupTime[VT][FL][CP])> = maxThreshold[VT][FL] **then**
2: **if** allocatedInstances[SID]<maxAllocate[SID] **then**
3: requestResource(FL);
4: newVID=scaleOut(FL;VT);
5: loadBalance(VID;newVID);
6: allocatedInstances[SID] + + ;
7: **end if**
8: **end if**

Fig. 56. Scale-Out algorithm for NFV [42]

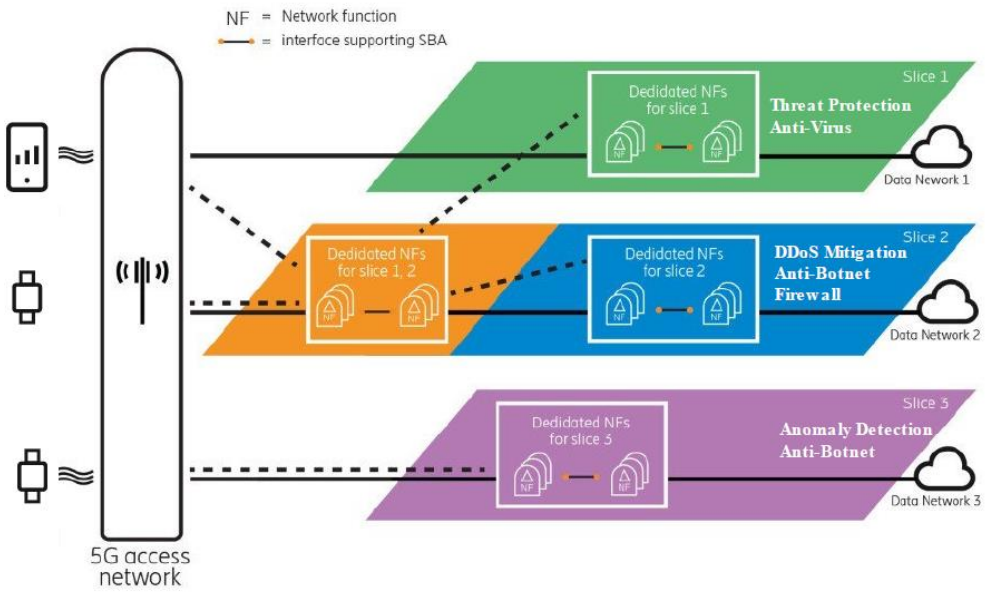


Fig. 57. SECaaS built upon Network Slicing [29]

13. Use Case

13.1 Rouge Base Station Detection

The Rouge Base station is a broad name for radio equipment and devices that impersonate a legitimate base station. Fake base stations are also known by other names like international mobile subscriber identity (IMSI) catcher, rouge, and stingray. They use radio interference between the mobile node and the base station, creating favorable conditions and tricking the mobile node into connecting to the rouge station instead of a legitimate station to attack the RAN. The rouge base station often mimics the frequency, cell ID, mobile country code (MMC), etc., as a legitimate base station and often transmits signals at higher power. The attacks performed can be characterized as active or passive, but one of the main attacks relates to passive eavesdropping on the user's identifiers over the radio interface. Other attacks are related to denial of service and camping to evade communications from legitimate base stations. Fraud attacks send spam SMS and advertisement to UE or create a public panic by sending false warnings. Rouge base can also jam signals and downgrade the UE to a lower RAT with fewer security mechanisms. Furthermore, it is cheaper and easier to set up a rogue base station with the advancement. Hence, we will further discuss how to detect such a base station and mitigate the problem at hand.

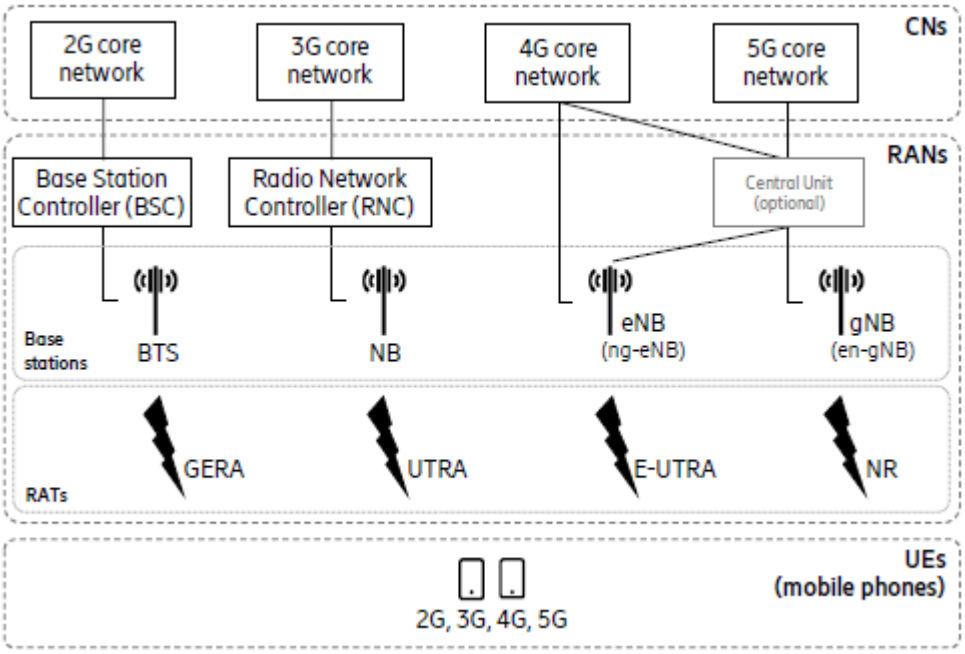


Fig. 58. An overview of mobile network. [45]

Radio access technology (RAT) with corresponding generations is shown in fig :

- 2G: GSM EDGE Radio Access (GERA)
- 3G : Universal Terrestrial Radio Access (UTRA)
- 4G: Evolved-UTRA (E-UTRA)
- 5G: New Radio (NR).

Note: 3GPP standards currently do not allow 5G to interwork with 2G for security reasons, i.e., a 5G network is restricted to measure down only to 4G and 3G, not 2G.

Only 4G networks can fully interwork with other generations, i.e., 4G can do the measurement on 5G,3G, and 2G.

13.1.1 Detection and Mitigation

The basic principle behind detection is to collect data and perform analysis of the radio environment measurements of the network and send that data to a server for analysis. Further can be categorized as:

- User equipment (UE) based
- Network-based system

13.2 A User Equipment (UE) Based

This approach will use a rule-based strategy that works well when the involved measurement parameters are known clearly. Here we determine if a rouge base station is present in the network from the view of UE. Then, the UE (mobile phone) can have a specialized application to analyze and collect the data from measurement reports. Further, this can send the data to the server on the internet for analysis.

13.2.1 CASE I

As we know, when UE power is on, it tries to attach to the strongest signal transmitted by a base station via attaching procedures and tests to authenticate each other. The base station initiates the identification procedure before authentication is performed between them. The base station sends an ID request to the UE, which response with a UE ID consisting of the IMSI or IMEI. The rouge base station will mimic the same procedure to get the UE ID and not respond with the authentication. In either case, it will fail at the authentication due to network security in RAT.

If the authentication fails, the UE can determine the base station as rouge and avoid camping on the site. In addition, the GPS coordinates, cell ID, EARFCN, and PLMN ID associated with the jamming station, and other information collected by the UE, can be recorded and compared to other rogue stations.

Once the UE is connected to a legitimate base station and complete the authentication mechanism, the UE will notify the network of the rouge base station by sending the stored information (GPS coordinates, cell ID, EARFCN, and PLMN ID) through signaling/message and updates the cloud server of the information.

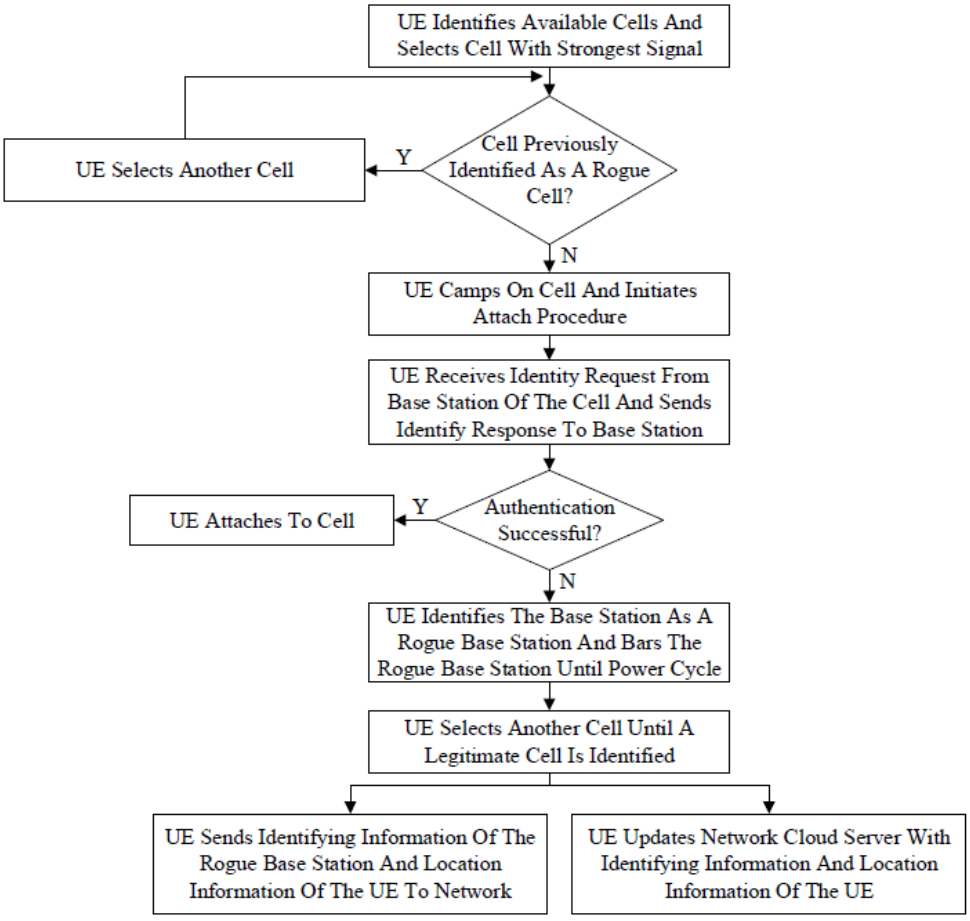


Fig. 59. CASE I [43]

13.2.2 CASE II

As we know, UE is usually a mobile device, and a legitimate base station is a fixed entity in the network. However, Rouge base stations can take advantage of this situation to move as if they were mounted on top of a car and capture as many UEs as possible by being mobile. To detect and mitigate such an approach, reverse mobility techniques are implemented.

In this approach, we can install Static UEs throughout the network in several locations and install an application that monitors the signal strength variation of detected cells. The parameter used to measure the variation in signal strength can be RSRP (Received Signal Received Power) or RSSI (Received Signal Strength Indicator) and RSRQ (Received Signal Received Quality) etc.

A well-planned legitimate base station is not expected to show wide variation in signal parameters as RSRP or RSRQ. In contrast, a mobile rouge base station can exhibit wide variation in the received signals. With received parameters, the application can run statistical analysis of signal characteristics using Machine Learning Algorithms.

The Static UE can determine information of rouge base station such as Cell ID, PLMN ID, MMC, etc., and detect if there is a sudden or unexpected variation of signal parameters indicating a rouge base station nearby. Also, this information can be sent with GPS location to the network operator with a peak time plot of signal variations indicating when the rouge base station was closest to UE.

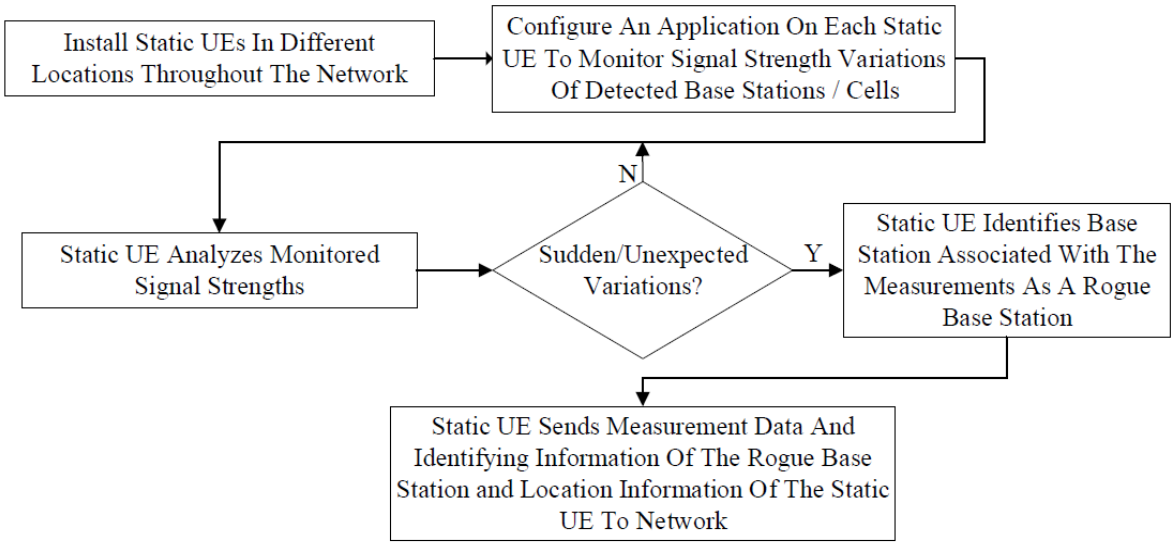


Fig. 60. CASE II [43]

13.2.3 CASE III

As we know, lower-order RATs are less secure such as 2G GERA and 3G UTRA, than 4G E-UTRA and 5G NR, and may also cause some backward compatibility issues when RATs are interchanged. Using this to its advantage, a rouge base station attempts to push UE to a lower order RAT. Therefore, in this case, our primary consideration will be downgrade attacks.

A rogue base station can typically capture more information from lesser secure lower-order RATs than a more secure higher-order RAT. When a UE tries to camp on a rogue base station, the rouge base station manipulates the cell reselection parameters such as Cell Reselection Priority and threshold to instruct UE to downgrade to a lower order RAT. Keep in mind that a legitimate base station will always assign a higher priority to a higher-order RAT (4G) than a lower-order RAT (2G).

Therefore, a UE can be configured to detect a situation in which a base station tries to reverse the priority and select a lower RAT. UE can detect and identify the base station as the rogue base station, avoiding camp on the cell until the next power cycle.

UE can store the information associated with the rouge station such as GPS coordinates, Cell ID, EARFCN (4G E-UTRA absolute radio frequency channel number), PLMN, etc., and look out for similar matching parameters to determine rouge base station. This information can be passed on to the network to further mitigate rouge station as in the case I.

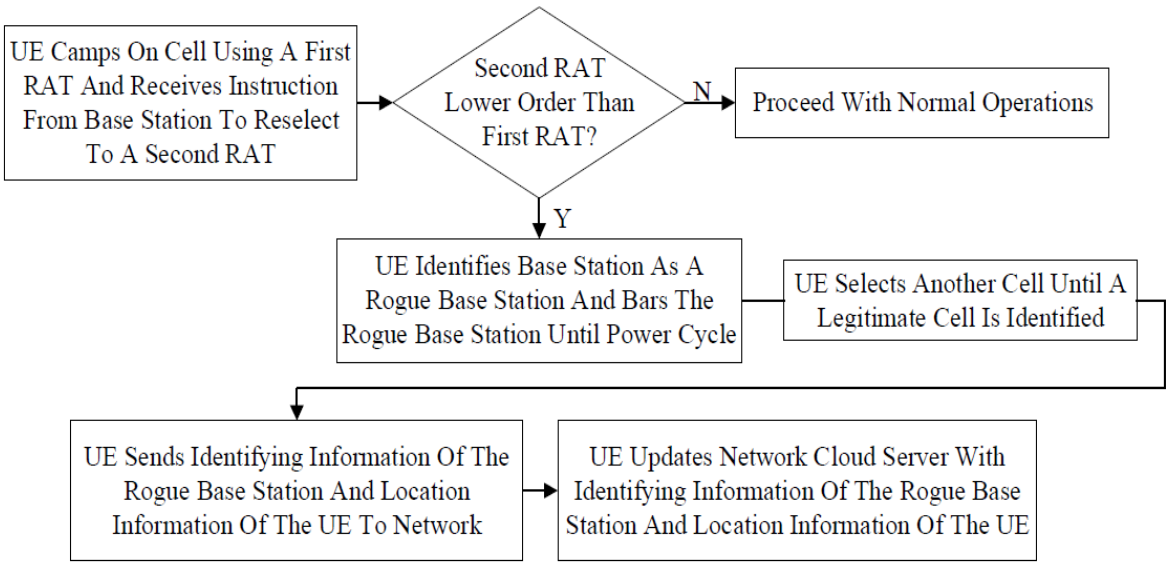


Fig. 61. CASE III [43]

13.2.4 CASE IV

In this case, detection and mitigation of rouge base station that jam signals are considered. Rouge base station may utilize a mechanism to jam signals and obstruct the signal between UE and Network by sending a signal with very high power. As a result, the rouge base station tries to connect with the UE because of a higher power, and legitimate base station signals are dropped due to less signal power.

Rouge base stations designed for such purpose are signal generators that transmit signals with a higher power in the same frequency as legitimate. A UE can be configured to detect abnormal high signal power (RSSI) and attempt to connect to a legitimate cell. The UE can switch to another radio access technology or bandwidth if it cannot find a legitimate cell.

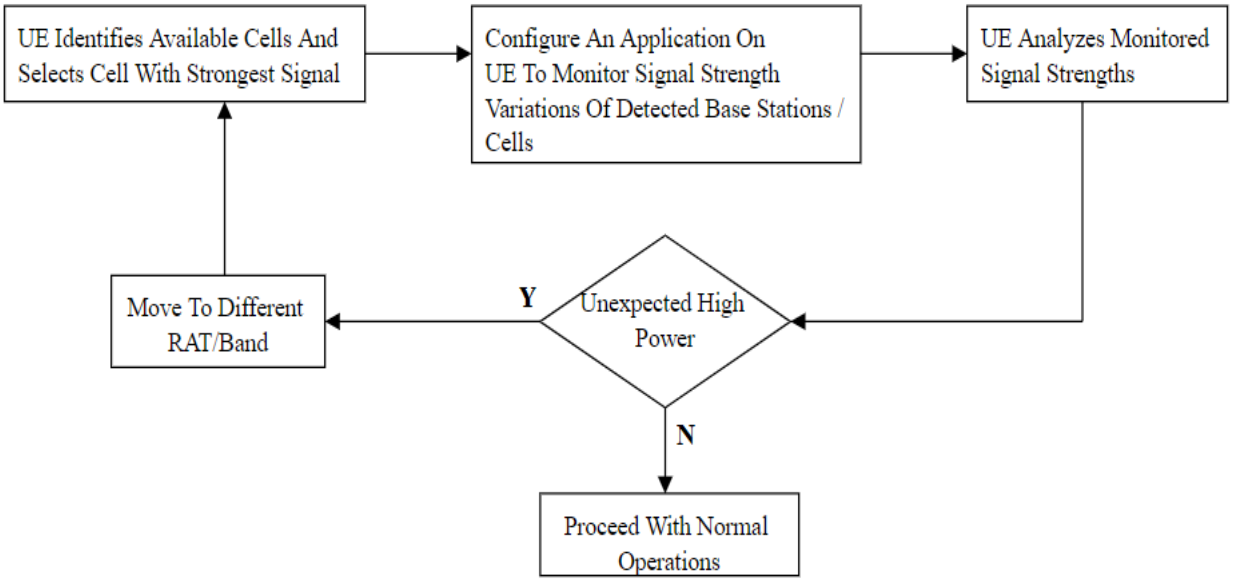


Fig. 62. CASE IV [43]

13.3 Network-Based System

This approach uses a rule-based strategy that works well when the relevant measurement parameters are known. Here we determine if a rouge base station is present in the network by considering a view from the Network, as a mobile network knows the global state of the system and local as well.

We can use a network functionality that uses the global state information of all the mobile devices (UE), network state, deployment, and configuration. By comparing the view of the mobile devices connected to the network and what the network intends the mobile devices to have, we can detect if a rogue base station is present. For example, if there is a discrepancy between the view of the cellular network and what the network expects, this may indicate that there are unauthorized stations. Keeping in view that the mobile device (UE) regularly updates the network with its local state information as a part of the normal operation of the network.

As UEs move in the area, they measure the signal strength of different base stations to connect and report the information to the network. However, it is difficult for the UE to determine whether any base station is legal or unauthorized from the collected information. On the other hand, the mobile network has the information about what base station will operate in what area, what signal strength to be expected, which identifiers will be exchanged, etc. Therefore, system tuning may be required depending on the parameter and threshold involved.

Deployment Scenario :

13.3.1 Step 1 - Data collection

UEs can be used as probes for data collection without any specialized software installed on them as they exchange measurement reports with the network. Although, depending on the transmission power, distance from UEs, the signal strength of legitimate base stations, and radio conditions, some UE may fall victim, being near a rouge station allows others to observe the false cell and report it to a legitimate base station.

Generally, a base station is the data collector that engages directly with UEs to receive reports using standard procedures. Other data collectors can be servers that manage RAN.

Data can be further classified into types :

Main Data: Measurement report. Identify the views of the network from the UEs' perspective

Auxiliary Data: Additional data to measurement reports like cell topology form the view of the network

UEs and RAN engage with each other on-air interface through standard Radio Resource Control (RRC) procedure enabling measurement report mechanism. The reporting mechanism is fundamental to all generations of mobile networks. It enables the network to decide what conditions are best suitable for an event, such as UE switching to a different base station. The RRC messages contain measurement configurations that include parameters like Cell ID, frequency and are termed as Measurement objects on which the UE is asked to perform the measurement. The reporting criteria and format are part of the reporting configurations. UE is triggered to submit a measurement report based on reporting parameters. The quantity, such as the number of cells, is specified by the reporting format. Both measurement object and reporting configuration are linked to providing a measurement identifier to identify the configuration. The measurement report may also contain parameters such as physical cell identifiers (PCI), received signal received power (RSRP), received signal received quality (RSRQ), signal-noise ratio (SNR).

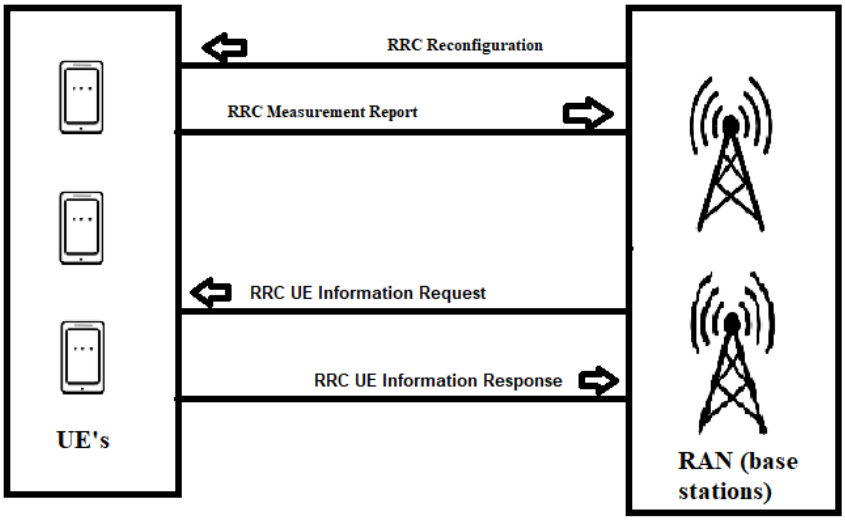


Fig. 63. RRC procedure

```
RRC {
   pdu value DL-DCCH-Message ::= {
    message c1 : rrcConnectionReconfiguration : {
      rrc-TransactionIdentifier 1,
      criticalExtensions c1:rrcConnectionReconfiguration_r8:{
          measConfig {
             measObjectToAddModList {
                 MeasObjectToAddMod {
                 measObjectId 4,
                 measObject measObjectEUTRA : {
                 carrierFreq 100,
                 allowedMeasBandwidth mbw6,
                 presenceAntennaPort1 FALSE,
                 neighCellConfig '10'B,
                 offsetFreq dB0}}},
       reportConfigToAddModList {
        ReportConfigToAddMod {
           reportConfigId 5,
           reportConfig reportConfigEUTRA : {
           triggerType event : {
             eventId eventA2 : {
               a3-Offset-6,
               reportOnLeave FALSE},
             hysteresis 2,
             timeToTrigger ms40},
          triggerQuantity rsrp,
          reportQuantity both,
          maxReportCells 4,
          reportInterval ms480,
          reportAmount r1}}},
      measIdToAddModList {
        MeasIdToAddMod {
                  measId 5,
           measObjectId 4,
          reportConfigId 5
}}}}}}
```

Fig. 64. 4G measurement report example. [44]

13.3.2 Step 2 – Analysis

One of the main functions of analyzers is to process the main data and the auxiliary data. Another function is data processing which prepares the data by parsing the measurement reports gathered from data collectors and auxiliary data. It can then apply different strategies, e.g., rules-based or machine learning, which uses the processed data and detects if any information indicates the presence of a rogue base station. If there are any deviations, they are flagged and termed as rouge.

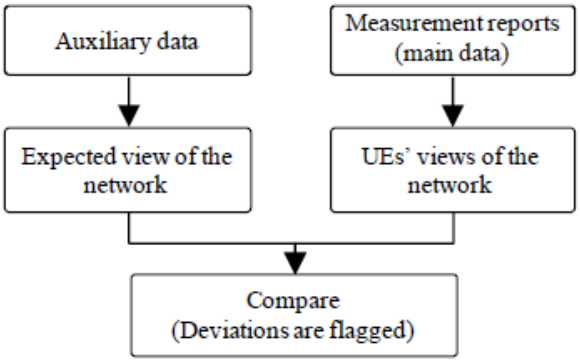


Fig. 65. Analysis step [45]

Example of Rules :

Rule: PCIs in range 0-50 → legitimate cell; otherwise → false cell

Rule: PCIs other than 25 reported between 18:00-8:00 → false cell

Rule: RSRP < -40 dBm → legitimate cell; otherwise → false cell

Rule: RSRQ > -7 dB → false cell; otherwise → legitimate cell

Fig. 66. Rules

The parameters, ranges, and threshold mentioned above can be hardcoded in rules or taken as input from the auxiliary data.

13.4 Effectiveness And Limitations

The approaches mentioned above that we have used to detect and mitigate rogue base stations are rule-based and categorized into User Equipment (UE) based and Network-based systems. The user equipment-based technique examines the network from the perspective of the user equipment to determine whether a rogue base station is present. However, it only has local information on the system. Additionally, they require specialized applications or servers and privileged root access to collect and analyze the measurement report, reducing the measurement for analysis. Therefore, this approach may give false positives as it does not know the view of the network, e.g., if a new base station is installed, all the UEs may determine it as false as never seen before.

On the other hand, a network-based system has the view perspective of the entire network and knows the global states of the system. As the approach relies on measurement reports from UEs, it functions if at least one or few UEs are connected to the legitimate network. If a false base station actively connects to all UEs so that no UE can send measurement reports to the legitimate network, this approach will not function.

Rule-based strategies work well if the relevant measurement parameters are known. So, it can be the first step in detecting rogue base stations before other strategies. However, there could also be cases when rules are impractical, and the values of one parameter fluctuate from minimum to maximum. In such cases, a more intelligent and robust strategy is required, like using Machine Learning (ML) algorithms and cloud-based services to have a complete proof detection mechanism easily implemented on UE and Network. Also, static rules are a disadvantage when the cell topology frequently changes, and manual maintenance is not practical in a large-scale deployment. Therefore, rules should automatically be updated according to cell topology. For such, we can use a real-time database.

14. Conclusion And Future Work

In conclusion, there are some important insights that we need to consider when starting future network deployment with security in mind. First, new services are evolving rapidly and becoming omnipresent. Not only do these applications consume large amounts of bandwidth, but they also have stringent delay and control requirements. Therefore, 5G networks need to be designed to be adaptable, resilient, and flexible to support these applications.

On the one hand, 3GPP defines a general security framework and key security measures based on higher-level cryptographic techniques such as NEA Encryption Algorithms and NIA Integrity Algorithms. On the other hand, however, the security requirements of some new services and the challenges posed by new technologies have not been adequately addressed.

5G technologies such as SDN and NFV are the basis for supporting services similar to 5G. In addition, these technologies will be used to enable network functionality and most security controls, highlighting the importance of a comprehensive security architecture. Therefore, it is essential to address the security challenges presented by SDN, NFV, and 5G.

This report highlights key security issues that could become more dangerous if not addressed adequately in 5G. I have also presented some security mechanisms and solutions for these problems. However, due to the limited independent (standalone deployment) and integrated implementation of these technologies in 5G, it is impossible to realize the threat vector fully. Similarly, as more user devices such as IoT are connected, and a new diverse set of services are provided, communication security and privacy issues will become more prominent.

Another problem is that the specific performance requirements of these devices may cause network operators to disable all possible security measures, for example, user plane encryption and integrity protection for better communication of these device classes. Finally, the virtualization concept presents new challenges for operators as it potentially creates new trust relationships between operators and third parties such as cloud service providers.

Additionally, it is essential to use standards, test suites, and proofs of concept across different security use cases to act as catalysts for implementation and realize the benefits of SDN, NFV, and 5G.

14.1 C-RAN

CRAN is a promising solution to common RAN challenges with distributed RRH and centralized BBU architecture, advanced multipoint transmit and receive technology, multi-standard support, virtualization technology on general-purpose processors, and services at the edge. As a result, CRAN can provide profitable growth to modern mobile operators with a competitive infrastructure to sustain in a dynamic market.

An overview of cloud radio access networks is provided to help you understand the basic system infrastructure deployed in 5G networks. Next, we discussed the major security issues that could pose a challenge to CRAN technology for 5G systems. To solve these challenges, I compared various requirements and solution properties. Since the installation is an integrated version of a single system, security threats cannot be fully detected at a single point in time. Similarly, if a large number of IoT-based user devices are granted access to services on 5G systems, the privacy challenges of CRAN information may become apparent in open forums. Therefore, addressing all new types of security threats and challenges is currently difficult and an open topic. Over the next few years, 5G technologies and services will be forced to introduce more components, creating some new challenges and solutions to address them.

14.2 Future Work

While it can be effective to detect cyber-attacks early and mitigate them quickly, it is advisable to stop them altogether with proactive measures. It can be achieved by applying AI/ML techniques for anomaly detection, enabling analysis of malicious agent behavior through deep packet inspection and traffic analysis, combined with past attacks. This approach can improve the detection and mitigation of attacks.

Maintaining confidentiality requires implementing anonymity-based techniques that mask the subscriber's true identity and substitute it with pseudo-identification. In this instance, encryption-based methods are also helpful. For example, communications can be encrypted before being sent to a service. Furthermore, to protect location privacy, we can use Obfuscation techniques like exclusive XOR cipher that will reduce the quality of location information. In addition, location masking algorithms such as cloaking are useful in combating some significant attacks on location privacy.

gNB exchanges on RRC level and NAS-level 5G cores require integrity protection from certain conditions. However, the level at which the UE implements this part of the specification and at least discards unprotected messages using level NIA1 is unknown. A UE that responds to an unprotected security mode command exposes the IMEI to a rogue network, thereby indirectly revealing the subscriber's identity. We may modify the 5G network's functional independent software radio implementation to disable integrity protection and use experimental SIM cards to ensure UE compliance. As far as privacy is concerned, enabling data encryption at the radio and NAS levels is entirely under the network operator's control. Therefore, it is necessary to check to what extent operators enable RRC, NAS, and user plane encryption. Operators should

deploy IPSec between all network functions on the network side. However, it is uncertain to what extent this will work. Auditing can only verify the strength of IPSec on the operator's network.

Collaborative efforts are possible when end-to-end security requires the development of highly coordinated and flexible standards involving multiple standards bodies. Additional standardization efforts may be needed to minimize gaps where 5G security standardization efforts are aligned, governed, and synchronized. It is also important to ensure that open-source software undergoes an appropriate review process and has adequate documentation. In addition, there should be a careful review of code in the public community. The government can also play a role in ensuring that security and privacy standards are strictly followed.

15. References

[1] N. U. M. O. C. O. a. A. J. A. Opeoluwa Tosin Eluwole, "From 1G to 5G, What Next?," IAENG International Journal of Computer Science, 45:3, IJCS_45_3_06, 28 August 2018.

[2] G. 01.02, "Digital cellular telecommunications system (Phase 2+), General description of a GSM Public Land Mobile Network," ETSI, 1996.

[3] M. Y. Rhee, Mobile Communication Systems and Security, John Wiley & Sons, 2009.

[4] P. C. Clint Smith, Wireless Networks: Design and Integration for LTE, EVDO, HSPA, and WiMAX, Third Edition, McGraw-Hill Education, 2014.

[5] Christopher Cox, An Introduction To Lte LTE, LTE-Advanced, Sae And 4g Mobile Communications, A John Wiley & Sons, Publication, 2012.

[6] P. Sharma, "Evolution Of Mobile Wireless Communication; Networks-1g To 5g," International Journal Of Computer Science And Mobile Computing, vol. 2, no. 8, p. 47 – 53, 2013.

[7] A. G. A. R. K. Jha, "A Survey Of 5g Network Architecture And Emerging Technologies," IEEE Access, vol. 3, pp. 1206-1232, 2015.

[8] ITU-R, "Recommendation ITU-R M.2083-0," ITU, 2017.

[9] H. K., S. P. H. J. Mamta Agiwal, "A Survey on 4G-5G Dual Connectivity: Road to 5G Implementation," IEEE Access, vol. 9, pp. 16193-16210, 2021.

[10] 3. T. 3. v. 15.3.0, "5G NR Physical Channels And Modulation," ETSI TS 138 211 V15.3.0, 2018.

[11] [Online]. Available: https://www.techplayon.com/understanding-basic-5g-nr-terminologies-subcarrier-spacing-frame-and-subframe-slot-and-ofdm-symbols/.

[12] 3. T. 2. v. 15.3.0, "System Architecture for the 5G System," ETSI TS 123 501 V15.3.0, 2018.

[13] M. J. N. P. M. P. Aman Sanwal, "Basis for Tracking Area Reconfiguration for 5G Networks," IEEE.

[14] I. N. Magazine, "historicjournals.itu.int," 2017. [Online]. Available: https://historicjournals.itu.int/viewer/3162/?return=1&css-name=include#page=1&viewer=picture&o=&n=0&q=.

[15] 3. T. 3. v. 1. Release, "5G Security architecture and procedures," ETSI TS 133 501 V15.2.0, 2018.

[16] P. Sahu, "5gblogs," [Online]. Available: http://5gblogs.com/5g-security-5g-aka-authentication/.

[17] W. L. D. P. D. A. M. G. B. V. L. Gerrit Holtrup, "5G System Security Analysis," arXivLabs, 2021.

[18] 3. T. 3. v. 15.8.0, "5G NR Overall Description Stage-2," ETSI TS 138 300 V15.8.0, 2020.

[19] B. D. M. D. a. J. M. M. Hadzialic, "Cloud-Ran: Innovative Radio Access Network Architecture," in Proceedings ELMAR-2013, 2013.

[20] H. L. C. Y. Y. L. S. Aleksandra Checko, "Cloud RAN for Mobile Networks—A Technology Overview," IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 405 - 426, 2014.

[21] "C-RAN: The Road towards green RAN," China mobile research institute, 2011.

[22] E. G. N. 2. V1.2.1, "Network Functions Virtualisation (NFV) Architectural Framework," ETSI, 2014.

[23] Y. W. C. H. F. D. N. H. X. Wenfeng Xia, "A Survey on Software-Defined Networking," IEEE COMMUNICATION SURVEYS & TUTORIALS, vol. 17, no. 1, pp. 27-51, 2015.

[24] ONF. [Online]. Available: www.opennetworking.org.

[25] J. Silver, "Cisco blog," [Online]. Available: https://blogs.cisco.com/ciscoit/sdn-101-what-it-is-why-it-matters-and-how-to-do-it.

[26] ONF, "SDN ARCH 1.0 06062014," Open Networking Foundation, 2014.

[27] E. J. S.-C. D. H.-M. F. A.-S. Juan Aranda, "5G networks: A review from the perspectives of architecture, business models, cybersecurity, and research developments," Novasinergia, 2021.

[28] R. P. Jover, 5g Protocol Vulnerabilities And Exploits, Bloomberg, 2020.

[29] "Security Considerations for the 5G Era," 5G Americas, 2020.

[30] D. J. a. C. Rajabhushanam, "Security Challenges and Solutions for Cloud Radio Access Networks," in IEEE, Nov. 2019.

[31] p. z. a. z. y. Feng yu tian, "A Survey on C-RAN Security," IEEE Access, vol. 5, pp. 13372-13386, 2017.

[32] 5. Americas, "The Evolution of Security in 5G," 2019.

[33] Y. Q. a. R. Q. H. Dongfeng Fang, "Security for 5G Mobile Wireless Networks," IEEE ACCESS, vol. 6, pp. 4850 - 4874, 2018.

[34] T. K. M. L. J. O. M. Y. a. A. G. I. Ahmad, "Overview of 5G Security Challenges and Solutions," IEEE Communications Standards Magazine, vol. 2, no. 1, pp. 36-43, 2018.

[35] T. K. M. L. J. O. M. Y. a. A. G. I. Ahmad, "5G security: Analysis of threats and solutions," IEEE Conference on Standards for Communications and Networking (CSCN), pp. 193-199, 2017.

[36] Y. W. W. Z. Shunliang Zhang, "Towards secure 5G networks: A Survey," Computer Networks, vol. 162, 2019.

[37] H. L. a. N. Kuntze, "Hypervisor-Based Attestation of Virtual Environments," 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress, pp. 333-340, 2016.

[38] J. B. a. K. Waldron, "5G And Zero Trust Networks," JSTOR, 2020.

[39] Cisa, "cisa.gov," 2021. [Online]. Available: https://www.cisa.gov/executive-order-improving-nations-cybersecurity.

[40] O. B. S. M. Scott Rose, "Zero Trust Architecture," NIST Special Publication 800-207, 2020.

[41] A. S. L. A. Jonathan Olsson, "5G zero trust – A ZERO-TRUST ARCHITECTURE FOR TELECOM," 2021.

[42] M. B. D. L. C. D. T. T. a. N. T. Yacine Khettab, "Virtual Security as a Service for 5G Verticals," IEEE Wireless Communications and Networking Conference, 2018.

[43] M. Venkata, "Rogue Base Station Detection Techniques," Technical Disclosure Commons, 2021.

[44] 3. T. 3. v. 1. Release, "Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC), Protocol Specification," ETSI, 2014.

[45] M. A. E. E. U. S. Prajwol Kumar Nakarmi, "Murat: Multi-RAT False Base Station Detector," Ericsson Research Security, 2021.

[46] "Security Orchestrator: Introducing a Security Orchestrator in the Context of the ETSI NFV Reference Architecture," 2015 IEEE Trustcom/BigDataSE/ISPA, pp. 1255-1260, 2015.

16. Acronyms

Term	Explanation
3GPP	3rd Generation Partnership Project
5GC	5G Core Network
AMPS	Advanced Mobile Phone System
NMT	Nordic Mobile Telephone
TACS	Total Access Communication System
GSM	Global System for Mobile Communication
SMS	Short Message Service
CDMA	Code Division Multiple Access
ME	Mobile Equipment
BSS	Base Station Subsystem
NSS	Network and Switching Subsystem
BTS	Base Transceiver System
BSC	Base Station Controller
MSC	Mobile Switching Center
SIM	Subscriber Identity Module
PSTN	Public switched Telecommunication Network
ISDN	Integrated Services Digital Networks
PLMN	Public Land Mobile Networks
VLR	Visitors Location Register
HLR	Home Location Register
AUC	Authentication Center
EIR	Equipment Identity Register
IMEI	International Mobile Equipment Identity
GPRS	General Packet Radio Services
EDGE	Enhanced Data Rates for GSM Evolution
IP	Internet Protocol
WCDMA	Wideband Code Division Multiple Access
UMTS	Universal Mobile Telecommunications Systems
HSUPA	High Speed Uplink Packet Access
HSDPA	High Speed downlink Packet Access
RAN	Radio Access Network
CN	Core Network
SGSN	Serving GPRS Supporting Node
GGSN	Gateway GPRS Supporting Node
LTE	Long Term Evolution Technology
WIMAX	Worldwide Interoperability for microwave access
OFDMA	Orthogonal Frequency Division Multiple Access
FDE	Frequency Division Equalization
MIMO	Multiple Input Multiple Output
UE	User Equipment
E-UTRAN	Evolved UMTC Terrestrial Radio Access Network
EPC	Evolved Package Control

HSS	Home Subscriber Server
MME	Mobility Management Entity
S-GW	Signaling Gateway
P-GW	Packet Data Network Gateway
PCRF	Policy and Charging Rule Function
eMBB	Enhanced Mobile Broadband
mMTC	Massive Machine Type Communication
uRLLC	Ultra-Reliable and Low-Latency communication
NE	Network Elements
DC	Data Center
SDN	Software Defined Network
NFV	Network Function Virtualization
IOT	Internet of Things
D2D	Device to Device
NFV MANO	NFV Management and Network Orchestration
SNR	Signal to Noise Ratio
BS	Base Station
MTC	Machine Type Communication
V2V	Vehicle to Vehicle
TDD	Time Division Duplexing
AUSF	Authentication Server function
AMF	Access and Mobility Function
DN	Data Network
UDSF	Unstructured Data Storage Function
NEF	Network Exposure Function
NRF	Network Repository Function
NSSF	Network Slice Selection Function
PCF	Policy Control Function
SMF	Session Management Function
UDM	Unified Data Management
UPF	User Plane Function
AF	Application Function
SA	Stand Alone
NSA	Non Stand Alone
DC	Dual Connectivity
MN	Master Node
SN	Secondary Node
DRB	Data Radio Bearers
SRB	Signaling Radio Bearers
SDAP	Service Data Adaptation Layer
PDCP	Packet Data Convergence Layer
URLLC	Ultra Reliable Low Latency
RLC	Radio Link Control Protocol
MAC	Medium Access Control
NAS	Non Access Stratum
RRC	Radio Resource Control

HARQ	Hybrid Automatic Repeat Request
PLMN	Public Land Mobile Network
SINR	Signalling to Interference plus Noise Ratio
RB	Resource Block
RE	Resource Elements
OPEX	Operational Expenditure
CAPEX	Capital Expenditure
HSPA	High Speed Packet Access
NTT	Nippon Telephone and Telegraph
FCC	Federal Communication Commission
FDMA	Frequency Division Multiplexing
MTSO	Mobile Telephone Switching Office
MMS	Message and Multimedia Service
TDMA	Time Division Multiple Access
ETSI	European Telecommunications Standards Institute
UTRAN	UMTC Terrestrial Radio Access Network
WMSC	Wideband CDMA Mobile Switching Centre
ATM	Asynchronous Transfer Mode
FDD	Frequency Division Duplexing
SC-FDMA	Single-carrier FDMA
IMT	International Mobile Telecommunications
NG-RAN	Next Generation RAN
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
PDU	Packet Data Unit
SDU	Service Data Unit
NAI	Network Access Identifier
IMSI	International Mobile Subscriber Identity
SUPI	Subscription Permanent Identifier
SUCI	Subscription Concealed Identifier
TMSI	Temporary Mobile Subscriber Identity
GUTI	Global Unique Temporary Identifier
GUAMI	Globally Unique AMF Identifier
NSSAI	Network Slice Selection Assistance Information
SST	Slice Service Type
SD	Slice Differentiator
NSSI	Network Slice Subnet Instance
VPLMN	Visited Public Land Mobile Network
HPLMN	Home Public Land Mobile Network
AKA	Authentication and Key Agreement
EAP-AKA	Extensible Authentication Protocol – Authentication and Key Agreement
ROHC	Robust Header Compression
TCO	Total Cost of Ownership
WCN	Wireless Cloud Network
RRH	Remote Radio Head
BBU	Baseband Unit

VNF	Virtual Network Function
NFVI	NFV Infrastructure
OSS	Operation Support Subsystem
MANO	Management and Orchestration
EMS	Element Management System
BSS	Business Support System
COTS	Commercial off-the-shelf
ONF	Open Networking Foundation
SS7	Signalling System No. 7
ACK	Acknowledgement
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
Ipsec	Internet Protocol Security
DoS	Denial of Service
PSK	Phase Shift Key
API	Application Programming Interface
CISA	Cybersecurity and Infrastructure Agency
PDP	Policy Decision Point
PEP	Policy Enforcement Point
SECaaS	Security As A Service
EARFCN	E-UTRA Absolute Radio Frequency Channel Number
RSRP	Received Signal Received Power
RSSI	Received Signal Strength Indicator
RSRQ	Received Signal Received Quality
SNR	Signal-Noise Ratio
BCCH	Broadcast Control Channel
PCCH	Paging Control Channel