Utilizing Enhanced VPN+ technology for achieving isolation of 5G network slices and improving security

Tariq Mahmood

Network and System Administration (NSA 2018-2020)
[60 credits ECTS]

Department of Informatics
The faculty of Mathematics and Natural Sciences

15-June-2020

Master Thesis

Tariq Mahmood,
University of Oslo
Oslo, Norway
E: tariqmah@ifi.uio.no
E: tarmahmood1@gmail.com

Mentor: Prof. Dr. Thanh van Do
Telenor Group, Telenor Research
Oslo Metropolitan University
Oslo, Norway
E: thanh-van.do@telenor.com

Table of Contents

Introduction

Mobile communications have made a tremendous and rapid growth over the last thirty years. All this growth has been achieved as a result of advances in micro-electric and software technology, innovative algorithms as well as continuous research and development in network technologies. These factors are prerequisites for the fast development of mobile communications which are shifting from first generations (1G) to second generation (2G), third generation (3G), fourth generation (4G) and rather soon to fifth generation (5G) communication. Industry standards such as 3GPP, 5GPP, GSMA, ITU, IETF has expectations to provide connectivity reaching 11.6 billion peoples with fifty times more connectivity within machine to machine and humans respectively from 2020 onwards while inheriting all the strengths of former generations of networks. Telecom operators, academia researchers, vertical industries and industrial partners are trying to bring improvements in infrastructure by innovating new and exploiting existing technologies to achieve the optimal results, which would satisfy the Customer's Quality of Service (QoS) as well as vertical industries in an efficient manner. Further universal connectivity will benefit to vertical industries and to personal users, but it will be attractive to malicious parties as well. They may wreak havoc from afar, and trigger financial, privacy, protection and national security disaster. The future development of mobile communication with augmented reality, autonomous, automation, machine-to-machine (M2M) communication entities, will face additional security challenges, such as new incoming attack vectors and machine-learning based exploits, as well as botnets and adversarial attacks.

1.1. Motivation

Tremendous growth in Mobile communications has provided higher data rates, optimal quality of service (QoS), less latency with a high capacity density, but it inflates the economic costs. This is a vast encumbrance on service providers in terms of capital expenditure CAPEX and operational expenditure OPEX. Further during the special events for example Olympic Games the mobile traffic is at peak in that particular region, high bandwidth is provided by increasing the CAPEX but during off days there is less traffic, which depicts that the operators have already spent the CAPEX along with OPEX still in running. Economically, it is not a feasible model for service providers. There is a need to have a network solution in which the services and resources

are scalable and flexible according to user/vertical industry requirements and have a full control in reducing CAPEX and OPEX as much as possible (Xiang, Zheng, & Shen, 2017). New vertical applications, such as smart home, smart cities, autonomous cars, e-health, augmented reality, put a new challenge on the mobile network and call for a change in paradigm of the implementation of the mobile network since the previous generations supported only human to human communication and in very limited extend Internet of Things IoTs (Zhang, 2018). Fourth generation mobile networks (4G – LTE) provides higher data rates, latency, less cost per bit and has a high-density capacity. We enjoy high-speed internet connectivity, streaming high definition videos (HDV), email etc., but it has some limitations too. Initially, it is a "one size fits all" scenario, and although it is possible to differentiate traffic in LTE through quality of service management or by prioritization of some application traffic the same amount of resources is allocated to each subscription. This limits the number of simultaneous connections and make LTE unable to support billions of heterogeneous IoT devices.

Fifth generation networks, which introduce substantial changes, with diverse services (smart home, smart cities, autonomous cars, machine critical applications). They support both varied services in terms of performance, auto-scaling, tailored Quality of Service, as well as support business in terms of CAPEX and OPEX. This diversification with new range of services supported in 5G does not correspond with the viability 4G LTE offered. Contention for all the resources among all the users of new service types and applications is strictly undesirable, and additionally, there is need to optimize network differently for meeting the performance requirements of these applications to run. For this reason, a concept of Network Slicing is decided and implemented (5GPPP, 2019).

Furthermore, the security and integrity of user data in 5G network will become a crucial standpoint for a successful deployment. With various services and applications, there will be numerous characteristics and requirements, i.e. in case of Internet of Things (IoT) related services, we require features in network that will guarantee latency, data rate, mobility, reliability, and security. To guarantee these requirements, it is necessary to provide adequate network performance, resource availability and a secure end-to-end isolation environment. Network slicing ensures these capabilities. With secure isolation, network slicing can prevent malicious traffic from third parties, thus strengthens the data integrity and privacy (Liyanage, Ahmad, Abro, Gurtov, & Ylianttila, 2018).

Many survey reports and white papers have attempted to examine specific requirements for network slice in 5G from the enterprise point view and point out technological constraints, but the same are not standardized yet. Nevertheless, they do agree that network slicing represents a central point of development in 5G (Ericsson, 2018; Foukas, Patounas, Elmokashfi, & Marina, 2017; Interdynamix Systems, 2018; Kazmi, Khan, H.Tran, & Hong, 2019) For that purpose, 3GPP has implemented a framework for network slicing and provisioning slice instances using orchestrators and automation.

Definition 1

According to 3GPP specification (3GPP-TS 23.501, 2019):

"A Network Slice is defined as a logical network that provides specific network capabilities and network characteristics".

"Network slices may differ for supported features and network functions optimizations. The operator may deploy multiple Network Slice instances delivering exactly the same features but for different groups of UEs, e.g. as they deliver a different committed service and/or because they may be dedicated to a customer".

"A single user equipment (UE) can simultaneously be served by one or more Network Slice instances via a 5G-AN (5G-Access Network). A single UE may be served by at most eight Network Slices at a time. The Access and Mobility Management Function (AMF) instance serving the UE logically belongs to each of the Network Slice instances serving the UE, i.e. this AMF instance is common to the Network Slice instances serving a UE".

Definition 2

IEEE Communications Magazine in May 2017 described network slicing as follows (Samdanis et al., 2017):

"Network slicing in 5G systems defines logical, self-contained networks that consist of a mixture of shared and dedicated resource instances, such as radio spectrum or network equipment, and virtual network functions."

Definition 3

IETF defined network slice as (Adrian Farrel, 2017):

"Network Slice – A Network slice is a managed group of subsets of resources, network functions / network virtual functions at the data, control, management/orchestration planes and services at a given time. Network slice is programmable and has the ability to expose its capabilities. The behavior of the network slice realized via network slice instance(s).

End-to-end Network Slice is a cross-domain network slice which may consist of access network (fixed or cellular), transport network, (mobile) core network and etc. End-to-end network slice can be customized according to the requirements of network slice tenants.

Network Slice Instance is an activated network slice. It is created based on network template. A set of managed run-time network functions, and resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the service instance(s). It provides the network characteristics that are required by a service instance. A network slice instance may also be shared across multiple service instances provided by the network operator."

According to above definitions, there is a conflict to define network slicing secure isolation, but one common aspect is the dissection of a physical network into multiple logical networks. Each slice should be managed independently (isolated), either with shared or dedicated resources, providing end-to-end solution. 3GPP has divided network slices into three segments (3GPP-TS 28.530, 2019):

- A segment of the Radio Access Network (RAN),
- A segment of the core network,
- A segment of transport network.

In this thesis, we examine the security aspects of the transport network layer between the Core network of 5G and the Centralized/Distributed units in the Radio Access Network (RAN). Provided that 5G core networks are softwareized and the same can be instantiated in clouds that are shared, thereby the multitenancy vulnerability (Brown et al. 2012) directly translates to the security of a core network of an operator. In order to remedy this situation, we propose an end-to-end solution by leveraging the VPN technology (IETF, 2016) as means for achieving harder isolation of network slices and provide substantial increase of security in the transport network segment.

Virtual Private networks (VPNs) have become most prominent services, provided by service providers (ISPs). They provide isolated connectivity to remote customers allowing computers to communicate within their network as though they are in same Local Area Network (LAN). This traditional VPN service was achieved by dedicated physical connectivity, which is obviously an expensive solution in terms of capital expenditure CAPEX. With improvement, VPNs were delivered over a shared infrastructure using multiplexing connectivity on the network, and with utilizing dedicated Packet Gateways (PGWs) as edge points for allowing end-to-end communication. This connectivity between edge gateways includes many techniques to transport private data by tunnels, which are implemented with IP-in-IP, Multiprotocol Label Switching (MPLS), segment Routing (SR), Layer-2 Tunneling Protocol (L2TP), Generic routing encapsulation GRE, Point-to-Point over Ethernet (PPPoE) protocols (described in section 1.5), etc., that is, in an overlay network, normally known as "soft isolation" of traffic. In a 5G network, vertical, industries, customers or users may request a connectivity with enhanced isolation from their services that has no effect on the throughput or latency of services provided to customer.

This thesis aims towards clarifying the concept of a hard isolation of network slice in a segment of a transport network, which will not affect the customer services regardless to the burst of traffic in a private network tunnel. For that purpose, various techniques such as Software Define Networking (SDN), Network abstraction, Network function virtualization (NFV) will be employed to engineer secure 5G network slices in a way that will provide performance assurance, flexibility, auto-scaling and modularity.

1.2. Problem statement

5G networks will allow customers to have a degree of control of network slicing in an overlay network, while operators will have a control on underlay network. Service isolation between the virtual networks is possible by soft slicing. They are easier to apply with IP-in-IP, Multi-Protocol Label Switching (MPLS), Layer-2 Tunneling Protocol (L2TP) on the overlay network, but this approach provides no guarantees in emergency use case scenario (with accumulated overhead in terms of latency and performance sacrifice for the URLLC (Ultra Reliable Low-Latency Communication) slice (Jain, Singh, & Babu, 2017; A. Kumar, Aswal, & Singh, 2013). If a slice of the tenant is supporting public broadband network services together with emergency services, a sudden volume of data or traffic bursts can lead to congestion of overlay VPN tunnels, which

will result to packet loss, delay and jitter. Furthermore, if the traffic between slices leaks despite an established policy, then the security can be compromised, and attackers will be able to traverse through other slices undetected. This can be ruinous to customer services and can influence vertical industry businesses that exploit that particular slice. Furthermore, in case of e-Health (Kapassa et al., 2019), it can accommodate the functionality of all three standard slices: eMMB, URLLC, mMTC (5GPPP, 2019) in one network slice.

eMMB (Enhanced Mobile Broadband): provides gigabytes of bandwidth, this includes ultra-high definition videos (UHDV), augmented reality,

mMTC (Massive machine type communication): deals with connection density that connects billions of sensors, smart buildings, smart home, automatic monitoring of health conditions,

URLCC (Ultra reliable low-latency communication): requires low-latency with high reliability for example remote surgery, augmented reality games.

For that purpose, the isolation of network slices should be performed on the transport segment between the core network and the Centralized/Distributed units in the same or another cloud/edge platform (see Figure 1). This becomes more prominent in cases where the Centralized unit shares the same cloud as the core network, attaining the same vulnerabilities of the multitenancy environment.

In this research work, we build a transport segment network slice between the Cloud Radio Access Network (C-RAN) and the cloud network. This shall provide complete separation within the underlying network fabrics in a way that the traffic from different services will utilize reserved compute resources with the help of Abstraction and control of Transport network (ACTN) framework that links to the SDN controller. The SDN controller provides the Virtual Network Functions (VNFs) and policies for accessing the 4G/5G cores in shared environments such as distributed clouds. Such architecture can solve some issues with roaming as well and provide additional security layer for network slices that use the ACTN framework (isolation, encryption, etc.).

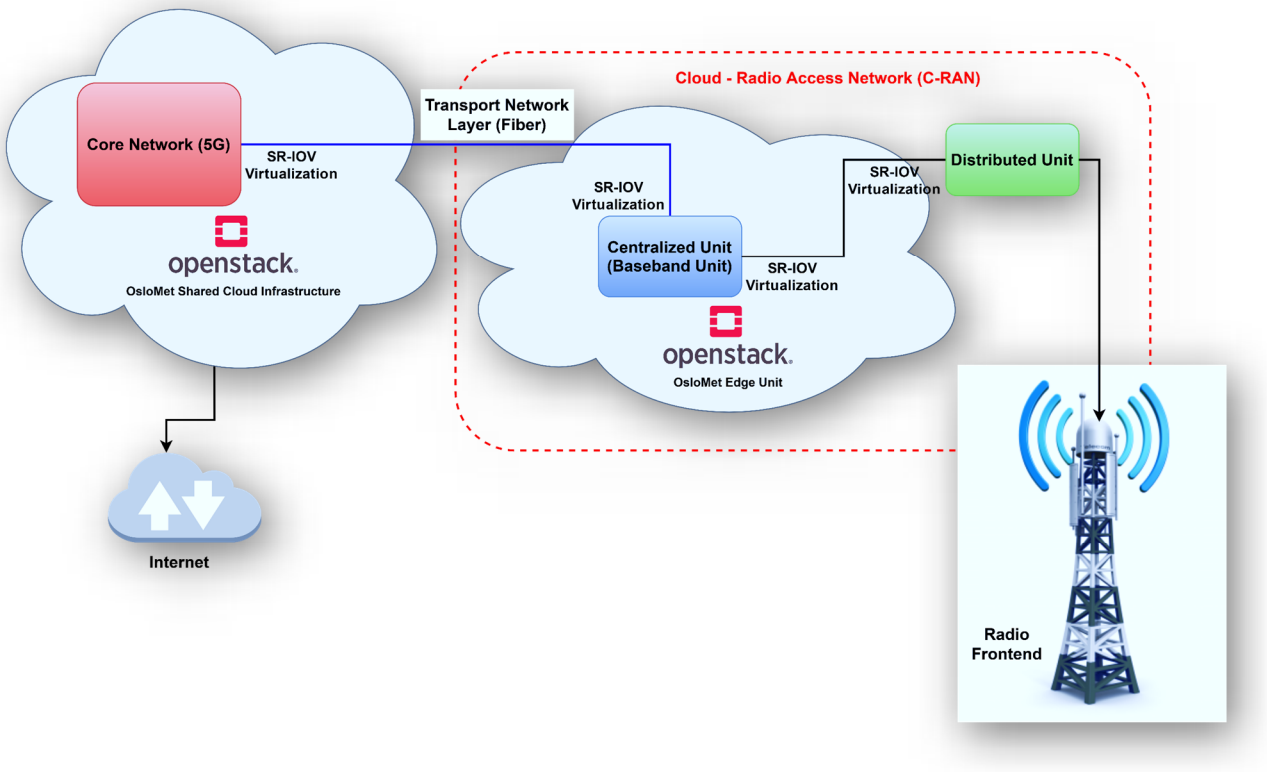

Figure 1. Enhanced VPN tunneling on the transport network (TN) between the 5GC (Core Network) and the Centralized Unit (Baseband Unit) in a shared or separate cloud

Additionally, the following sub-problems shall be examined:

Sub-problem 1: Mechanism to achieve **integration** between reserved or dedicated VTN (virtual transport network) underlay resources to overlay VPN.

- Integration needs to be implemented in a flexible and scalable manner so that it can be widely used in an operator's network to support eVPNs customers.

Sub-problem 2: Abstraction topology in applying policy to extract information that may represent the ability to connect across the network. It should present a connectivity graph that is independent of underlying technologies, as well as topology that can be used deliver network services uninterruptedly. This can be achieved by using the ACTN framework.

Sub-problem 3: Network slice isolation and enhancing security on the transport layer

By utilizing encryption in the tunnel, the VPN network can help strengthen the isolation between network slices when needed. This however adds a certain amount of overhead and impacts the performance of the network in general. The overhead will be also benchmarked.

1.3 Methodology

To examine the described problematics, we utilize the design-science methodology to set up an environment as described in Figure 1, where two endpoints of the 5G network are established: the Core Network (5GC) and the Centralized Unit (CU). The Core network is the system that is responsible for all the 5G-related networking with regard routing in the mobile network, authenticating users (devices, IoT entities), network slicing functions, roaming, user management etc. The other system is the Centralized Unit, also known as Baseband Unit, which performs operations such as Radio Resource Control, Fast Fourier Transform/Inverse Fast Fourier Transform, translating physical messages from the radio frontend into higher layer communication protocol messaging etc. This is a concept stemming from functional splitting of the radio frontend, for the sake of offloading the functionality of the base station and move a split portion of it into the cloud. This split distribution we refer to as "Centralized Unit", which is a hallmark for network slicing in 5G, to which the radio frontend connects with Remote Radio Units (also known as Remote Radio Heads) hosting the antennas from which the mobile user equipment is receiving connection. The link between the Core Network and the centralized unit is direct in a perfect case scenario, but in many instances that will not be possible. For best performance, we show that the direct connection using fiber links will yield a full link capacity, allowing for optimal 5G networking on top.

Following the instituting of these two endpoints, we utilize the best practices from virtualization to deploy VNFs (Virtual Network Functions) in the cloud datacenter, using generic server hardware. The hardware-level virtualization of the network functions is performed using Intel's SR-IOV drivers, which are mapped into Docker containers for allowing the container-level virtualization to directly communicate with the underlying hardware. This minimizes additional overhead incurred by the virtualization and allows for more granular control.

As a final step, we instantiate an enhanced VPN tunnel between the virtualization endpoints and we perform an evaluation. The evaluation consists of measuring bandwidth at the hardware level, virtualization level and hardware-VPN level as well as virtualized VPN level to determine the level of impact the encryption has on the 5G network. In parallel, the 5G communication is instantiated between the Core Network and the centralized Unit and the performance degradation is evaluated. To conclude, we provide a clear understanding of the repercussions a Software-

Defined VPN has on the transport network between the network core and the Centralized Unit, due to the encumberment of the hardware the 5G infrastructure is running on.

1.4 Organization of the thesis

The master thesis is organized as follows:

- Chapter 1: Introduction about the mobile communication development and motivation of this thesis work, as well as a problem statement along with used methodology that has been elucidated comprehensively.
- Chapter 2: Background provides detailed knowledge about the relevant communication technologies, security of mobile networks, networking slicing, virtual private networks and OpenAirInterface5G that are necessary to be comprehended for the understanding of this thesis.
- Chapter 3: An implementation part, in which the Realization of Encrypted Transport network slice and instantiating an enhanced Software-Defined VPN tunnel between core network's SDN controller and Centralized unit in the cloud are explicated in detail.
- Chapter 4: Evaluation section, in which an uninformed testing of the implemented environment is conducted, in order to determine the impact the Software-Defined VPN tunneling encryption can have on the transport layer in best case.
- Chapter 5: Discussion chapter providing a short notion of the lessons learned, as well as some suggestions for further improvement, and
- Chapter 6: Conclusion, for concluding this work and decide whether this approach possesses the prospects of utilization in the real world.

1. Background

1.1. GSM

First generation cellular networks were launched in mid-1980 and were based on analog technology. They were supporting only voice traffic. Small lightweight cordless handsets were being introduced for the communication up to 100 meters, attached to the telephone network at home, In Japan they were referred to as “Personal Handy Telephone” (PHS) (Figure 2). Similar work was ongoing in the European region by the name of “Telepoints”, secondly along with handsets paging technology was also increasing rapidly in the mid of 1980’s which was cost effective solution compared with the handset communication. A paging alert was sent to the particular individual and then they would establish a contact through telephone (Gsma, 2019b; Terasvuo, 2019).

Figure 2. Personal handy Telephone (Linge, 2019)

The first generation (1G) networks used analog signals with speeds up to 2.4 Kbps. Nippon Telephone and Telegraph (NTT) had introduced in Japan the first cellular system, whereas Nordic Mobile Telephone (NMT) and Total Access Communication System (TACS) were popular and introduced in the European region (D. S. Kumar & Meraj ud in Mir, 2015). United States (US) launched first mobile system called Advanced Mobile Phone System (APMS) as 1G mobile system, which allowed voice calls within the country but with some limitations:

- Poor voice quality,
- Large phone size,
- No security,

- Very slow speeds,

Mobile phones were dependent on operators, they were property of service providers (D. S. Kumar & Meraj ud in Mir, 2015).

As the economies of the world were growing, the concept of globalization was emerging onto global stage. Emigration was on the rise all over the world due to globalization and economists were expecting the growing demand of communication in parallel. The European market, especially the Scandinavian countries, showed interest in expanding the cellular network to expand the communication to other countries (Terasvuo, 2019). All the developing countries started to mature their own telephones to larger volumes. CT2 standards (Chen & Zhang, 2004) were introduced by United Kingdom (UK) to be followed by the development of cordless telephones (Terasvuo, 2019). Both APMS (D. S. Kumar & Meraj ud in Mir, 2015) and TACS (Linge, 2019) were built on the basis of analog technology. However, due to increase in users, the scalability issues transpired, which pointed towards a need for efficient network technology that can handle the desired scalability and flexibility. Conference of Postal and Telecommunication (CEPT) was an instigated committee for developing the standard for digital communication in 1982 (Gsma, 2019b) This committee ensured the setup of a new system, which could handle international voice call support, quality of service, service cost and support for new Integrated Services Digital Network (ISDN). GSM was officially started in 1992 and there was a dramatic increase in users, namely more than one million subscribers. In 1993, due to recognition and as an efficient standard, GSM was established also outside of Europe, such as the USA, United Arab Emirates, South Africa, Iran, Syria and New Zealand.

1.1.1. GSM Architecture

GSM is a well-acknowledged standard for cellular communication. 2G architecture is based on several active components and subsystems as shown in Figure 3.

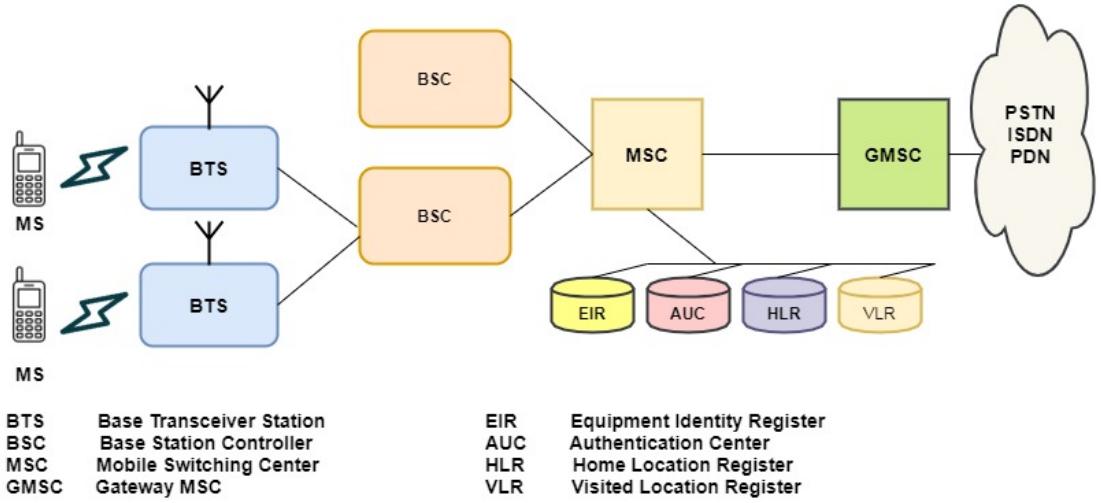

Figure 3. GSM architecture. (Eberspacher, Vogel, & Bettstetter, 2001)

The cells are connected and communicate through radio transceiver stations called *Base transceiver stations* (BTS) (Ali, Faisal, Rahman, & Haq, 2017). All the BTS are connected to or controlled by *Base stations controller* (BSC), then the traffic is routed through *Mobile switching cente*r (MSC) that is being received from combined BTS. GSM architecture is divided into three main sections as below shown in Figure 4 (Eberspacher et al., 2001; Sempere, 2002).

- Mobile stations (MS),
- Base station subsystem,
- Network subsystem.

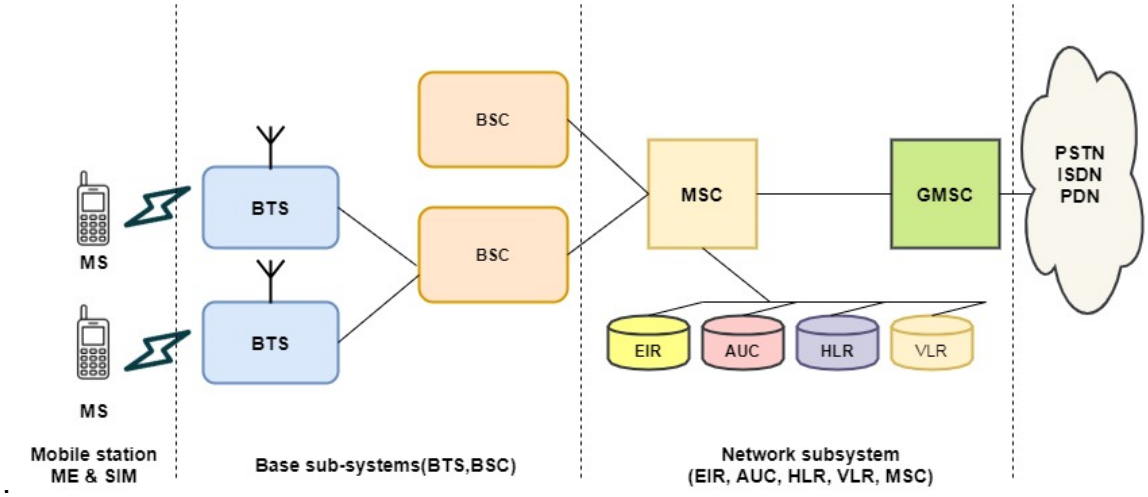

Figure 4. Components of Mobile and Base Stations subsystem (Eberspacher et al., 2001)

1.1.2. Mobile Stations

a) Subscriber Identity Module

Mobile stations (MS) consist of mobile equipment's (phone or terminals) and Subscriber identity module (SIM) (Sempere, 2002). The SIM gives the personal information of a subscriber to the system. It contains the international mobile services subscriber identity (IMSI) (Gsma, 2019c). Additionally, the SIM has multiple features that include the receiving and making of calls across the different, or same mobile network stations (Eberspacher et al., 2001). One supplementary feature is that it can store the list of subscribers and mobile network locations for fast network convergence, in the Home Location Register (HLR). The IMSI consists of three parts:

1) Mobile country code (MCC): It is used to identify the country network. Each country has its unique MCC,
2) Mobile code network (MCN): This is a part of IMSI to separate the different networks within the same region of country. In Norway there are many operators. For identification, i.e. separate MCN for Mycall, Telenor and Alika mobile operators are used,
3) Mobile subscriber identification number (MSIN): The third portion of the IMSI is MSIN identifier, which is the ID of subscriber (Eberspacher et al., 2001).

b) Mobile Equipment (ME)

Mobile Equipment is a category of MS characterized by radio and service function where a radio transceiver is used to communicate with the base station system. GSM 900 and GSM1800 are two frequencies used in this stations to communicate with Base station system (BSS) (Das, 2010), as shown in Figure 4.

1.1.3. Base Station Subsystems

a) Base Transceiver Station

Represented in figure 3 is the *Base Transceiver Station*, a part of the Base-station subsystem category. This provides the radio frequency channels to connect with the subscribers (the *Mobile Stations*), The interface between MS and BTS is called air interface or um interface (Eberspacher

et al., 2001) with associated protocols. While A-bis connection (interface provides communication between BTS and BSC using protocols) is between BTS and *Base station Controller* (BSC) (Heine, 1999). The BTS main functions are:

1) Transceiver TRX,
2) Combiner,
3) Control function,
4) Clock module,
5) Power amplifier,
6) Multiplexer,
7) Operation and Maintenance Module.

b) Base station Controllers

Base stations controllers BSC are the component of BSS. They control the multiple BTSs. All the BTSs are connected through A-bis interface (Eberspacher et al., 2001; Heine, 1999). BSC provides the channels such as signaling and traffic, this traffic is then used by MS to communicate with other subscribers within the BSS (3GPP-TS 23.002, 2001). In cellular communication, the analog voice is converted to 13 kbps digitized form by converter and this same information is converted to 64kbps by BSC (Palanivelu & Nakkeeran, 2008). For that purpose, the BSC is characterized with the following functions:

1) BSC reserves radio frequencies and manages the handover between multiple or connected BTSs using an A-bis interface (ETSI-GSM 08.52, 1996),
2) Controlling and routing the signal to Manager System control (MSC) through A interface (used to link between BSC and MSC. A interface is a 2-Mbps standard Consultative Committee on Telephone and Telegraph (CCITT) digital connection (3GPP-TS 23.002, 2001; ETSI-GSM 08.52, 1996; Garg & Wilkes, 1998; Palanivelu & Nakkeeran, 2008),
3) It provides authentication, encryption and decryption of data (ETSI-GSM 08.52, 1996).

1.1.4. Network Subsystem

a) Mobile Switching Center

The *Mobile switching Center* (MSC) receives the signals from BSC through an A-bis interface (ETSI-GSM 08.52, 1996). Functions performed by MSC are given below (Eberspacher et al., 2001; Nokia Networks Oy, 2001):

1) Signal routing,
2) Routing path, including fix network switching Nodes,
3) Service feature processing.

Apart from this, its function is to allocate and administer the radio resources and subscribers. MSC has four main elements represented in Figure 5.

i. Home Location Register (HLR)

The HLR provides the administrative information of the subscribers in particular GSM network (Eberspacher et al., 2001). The information whether the subscriber is eligible for the service or not is obtained from here. Moreover, it contains the mobile ISDN network information and the current location of the mobile station (Palanivelu & Nakkeeran, 2008). In other words, the HLR represents a database comprised of administrative attributes about a subscriber in MSC. It is a permanent database registry (ETSI-GSM 08.52, 1996).

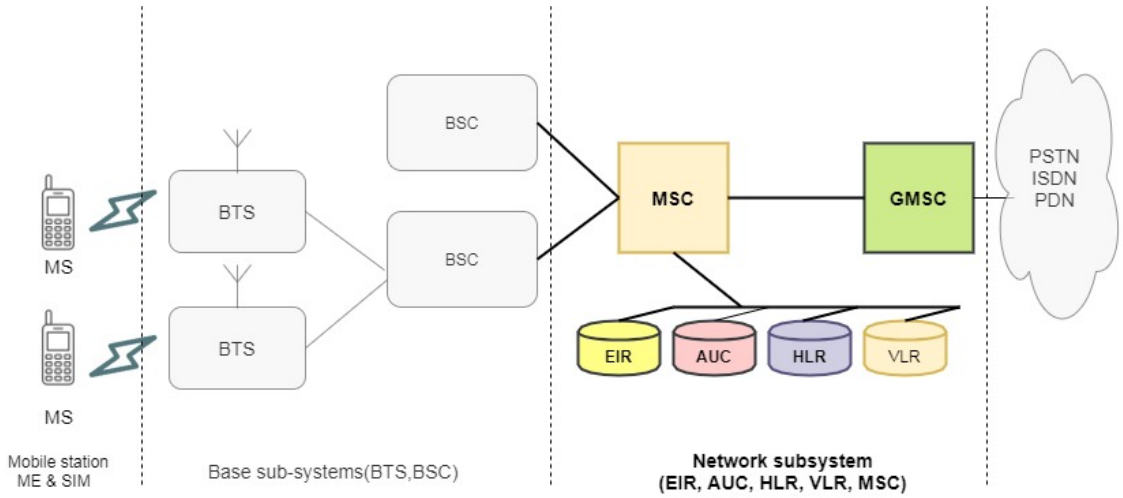

Figure 5. Network Subsystem elements (Eberspacher et al., 2001)

ii. Visitor Location Register (VLR)

Visitor Location Register is one of the elements of the network switching subsystem in MSC. It contains information of subscribers that are in the service region of MSC (Terasvuo, 2019). This information is registered in VLR and if there is no update in HLR, it is stored there as well. This implies that when a mobile station comes to new MSC area, the VLR registers the subscriber of new mobile station in order to use the services of the network. Unlike the HLR, a VLR database is temporary (Nokia Networks Oy, 2001).

iii. Authentication Center (AUC)

Authentication of subscribers takes place in the AUC element of the MSC, for the mobile stations that want to use the network. AUC relates to HLR, which gives the information of authentication parameters and ciphering keys to ensure security of the operator network. A H-interface(provides transmission of short messages between MSC and SMS-Gateway using MAP/H protocol) is used between AUC and HLR for communication (3GPP-TS 23.002, 2001).

Authentication parameters are following:

1) Secret key Ki. This same key is on the SIM,
2) Encryption algorithms A3 and A8, also on the SIM.

This key Ki, together with the algorithms A3 and A8, form an authentication triplet. VLR saves this information and corresponds to MSC. Authentication takes place in the MSC. This encryption is used to authenticate the MS-BTS link (Anderson, 2008).

iv. Equipment Identity Register (EIR)

Equipment Identity Register stores the information of International Mobile equipment Identity (IMEI) (ETSI-GSM 02.16, 2000). When a MS is connected, the MSC checks the IMSI from HLR and IMEI from EIR. EIR consists of three lists (ETSI-GSM 08.52, 1996):

1) White list, which allows MS to operate normally,
2) Grey list is used to find the defect in the MS,
3) Blacklist uses the information of stolen IMEI and thus are not allowed to operate in the network.

1.1.5. GSM interfaces

In GSM network architecture, there are many interfaces that are used to connect the elements within the network containing control channels and radio frequencies information, as shown in Figure 6 (Eberspacher et al., 2001). The Table 1. GSM interfaces and their function in the core networkTable 1 summarizes the functions of each interface.

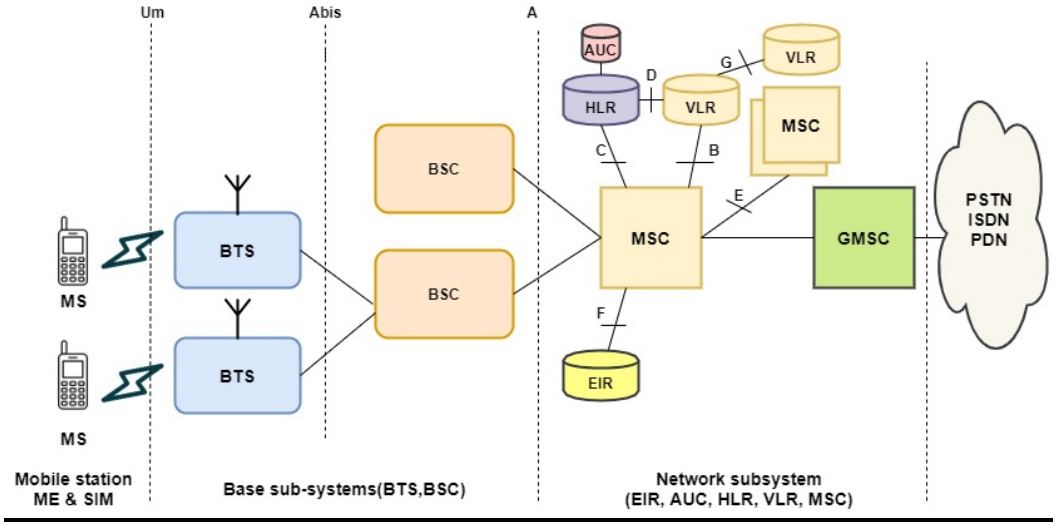

Figure 6. GSM interfaces (Eberspacher et al., 2001)

Table 1. GSM interfaces and their function in the core network

Interface	Description
Um	Air interface that connects between MS and BTS, which carries data and control information. (ETSI-GSM 02.16, 2000)
A-bis	This interface is used to link BTS and BTS. Radio frequency and radio equipment in BTS are controlled on this interface. (3GPP-TS 23.002, 2001)
A	It is used to link between BSC and MSC. A interface is a 2-Mbps standard Consultative Committee on Telephone and Telegraph (CCITT) digital connection. (Garg & Wilkes, 1998)

B	It is used for information sharing between MSC and VLR (see section 2.5)
C	It is used between HLR and MSC (see section 2.5)
D	Communication between HLR and VLR.
E	Link between MSC and MSC, Handover to another cell or network in order to connect with destination MS. (ETSI-GSM 02.16, 2000)
F	Linking MSC and EIR. This link is used to exchange the IMEI information of MS to get the access for network. (Devra & Sharma, 2016)
G	Connection between to VLR of different MSCs to communicate the information of subscriber (see section 2.5)

1.1.6. GSM signaling protocols

GSM has a layered model that is used to link the communication between the two end systems. The model works with the same concept of information passing to upper layers and vice versa, which ensures the information is transmitted and received. As shown in Figure 7, the GSM architecture is divided into three layers (Rahnema, 1993). Layer-1 uses the channel structures linking to BTS over Um interface. The channel structures of GSM are logical as well as physical. Physical channels are specified by time slot or carrier radio frequency (RF). GSM used Time-Division Multiple access and Frequency-Division Multiple access (TDMA/FDMA), namely a FDMA of 25 MHz bandwidth divided into 124 carrier frequencies of 200 KHz bandwidth each. BS are assigned by these RF (Sempere, 2002). For two-way voice communication, it is required at least one radio channel between BTS and MS, in which one end becomes the forward uplink and other end becomes the downlink or reverse channel (Harte, Bromley, & Davis, 2008). GSM uses the TDMA technique by dividing 8 voice calls, or eight bursts are grouped into one channel. Each burst of TDMA channel is sent and received on different frequencies (Eberspacher et al., 2001).

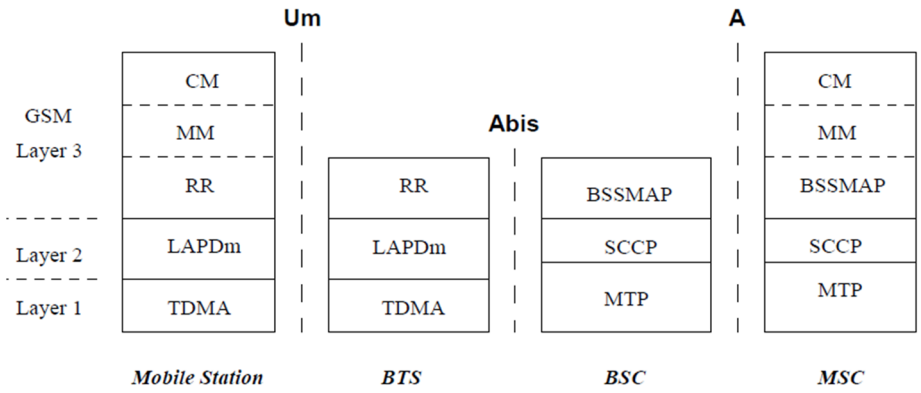

Figure 7. GSM signaling protocols (Eberspacher et al., 2001)

On Layer-1, the GSM architecture defines the logical channel that consists of two channels, which are traffic channel and signaling. Traffic channels are used to carry information such as speech, fax and data. Layer-3 information is not carried by traffic channel. As shown in Figure 8, these are channels that are combined to map physical and logical channels in GSM, that is, which BTS will adopt the selection of time slots, accordingly.

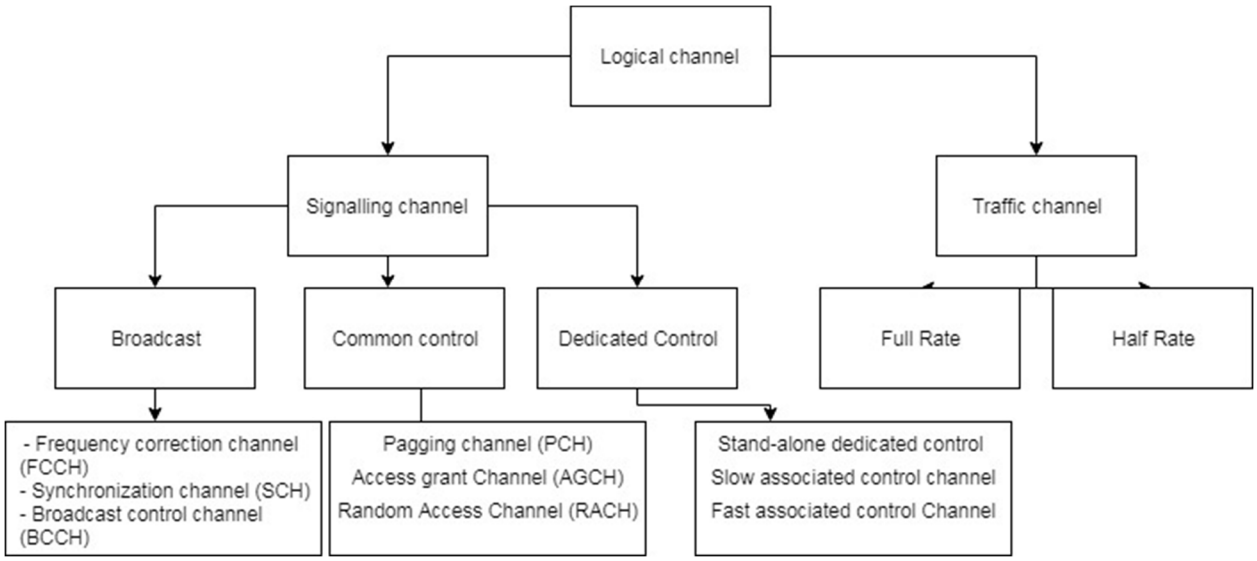

Figure 8. Channel Structure in GSM (RF Wireless World, 2012)

Broadcast channels are point-to-multipoint from BTS to MS. Frequency Correction Channel (FCCH[1]), as shown in Figure 8, is transmitted by BTS to MS. The MS tunes to exact RF of BTS

1 FCCH performs frequency correction/synchronization of MS in every 10 sec transmission of frequency bursts.

cell. The Synchronization Channel (SCH[2]) consists of transceiver code of BTS identity code (BTSIC), which helps MS to adjust accordingly on TDMA in GSM network. The Broadcast Control Channel (BCCH[3]) consists information that is being compared with the SIM services, to establish connection in the network. *"Common control"* is another category of signaling channels in GSM network, which are also point-to-multipoint downlink. The Paging control channel (PCH[4]**)** is part of common control used to alert the MS end system for establishing the connection**. "***Dedicated control"* are point-to-point channels. They are used for call setup and handle the requirements of handovers in the network (ETSI-GSM 05.01, 1992).

Layer-2 is data link layer. In Integrated Services Digital Network (ISDN), LAPD protocol (performs reliable transmission information over A-bis interface) is used, whereas the LAPD is modified for MS and named as LAPDm. This layer carries the frames of FCCH, BCCH, SCHand Access Grant channel (AGCH[5]). LAPDm makes the link reliable at BTS with CRC and ARP requests (Eberspacher et al., 2001).

1.1.7. Strengths and weaknesses of GSM

Strengths:

- Digital transmission,
- Transmission is encrypted, which is secured.

Weakness:

- Supports only voice, does not provide video,
- Only SMS service, which is limited data,
- It is a one-way authentication.

1.2. 3G (3rd Generation Technology)

The standard of mobile communication was indeed GSM, but due to advancements in technical and economic causes, GSM is followed by Universal Mobile Telecommunication

2 SCH performs synchronization between MS and BTS in every 10 frames of synchronization bursts.
3 BCCH is used to control broadcast information (carrier frequencies of BTS, parameters) within cell.
4 PCH checks periodically if a network required connection to MS (incoming call/short message).
5 AGCH performs physical channel allocation (timeslot) of UE.

Systems (UMTS), also known as 3rd Generation mobile system – 3G (Eberspacher et al., 2001). The 3G standardization was being developed by several organizations, including the ETSI Special Mobile Group (SMG). The European Commission was funding several research programs, such as Advance Communication Technologies and Services (RACE I & ACTS). Both ETSI and ARIB selected WCDMA as their radio interface, which forced the telecommunications industries to join the program 3GPP whose goal was to work on a process that produces 3G systems based on Universal Terrestrial Radio Access (UTRA) radio interface along with enhanced GSM/GPRS core network (Korhonen, 2003).

Figure 9. 3G mobile phones (gsmareana, 2019)

1.2.1.3G Architecture

Universal Mobile Telecommunication System (UMTS) is the next generation technology succeeding the 2G – GSM and GPRS. UMTS architecture consists of three domains (Kaaranen, Ahtiainen, Laitinen, Naghian, & Niemi, 2005), as shown in Figure 10:

- User Equipment (UE) (Kaaranen et al., 2005),
- UMTS terrestrial radio network (UTRAN) (Kaaranen et al., 2005),
- Core network (CN) (Kaaranen et al., 2005).

For the purpose of clarity, the interfaces are exempted in the figure.

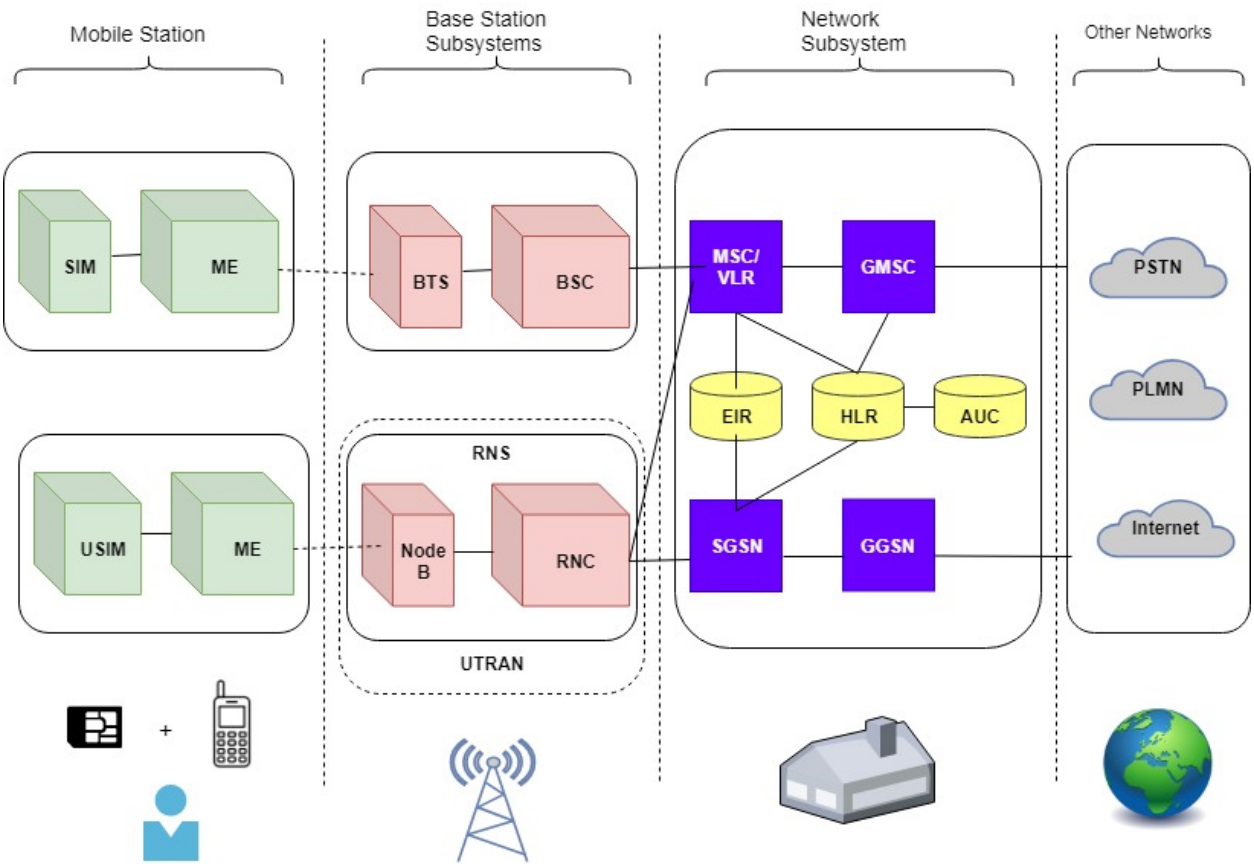

Figure 10. UMTS 3G Architecture

a) UMTS Terrestrial Radio Network

UMTS Terrestrial Radio Network (UTRAN) consists of Ba*se Stations* and *Base Station Controllers* (Kaaranen et al., 2005). A Base Station (BS) in UTRAN is referred to as Node-B and its controller is called *"Radio Network Controller"* (RNC), (see Figure 10).

The Node-B is the physical unit for transmission and reception of radio signals with cells and is usually mounted on building roofs and specially protected housing units close to radio towers. BTS serves as cell. The radio interface called Uu is connected to UE which is used for conversation of data to and from Uu interfaces, therefore the main functions of Node-B are as follows (Karim & Sarraf, 2002):

- Air interface TX/RX,
- Modulation /Demodulation.

The RNC is connected to one or more Base Transceiver Stations (BTS). Functions of RNC are (Kaaranen et al., 2005):

- Radio resource control,
- Channel Allocation,
- Power control settings,
- Handover control,
- Ciphering,
- Segmentation and Reassembly.

On the other side, RNC are connected to Core network (CN) over Iu interface. See Figure 11 UTRAN overview with interfaces.

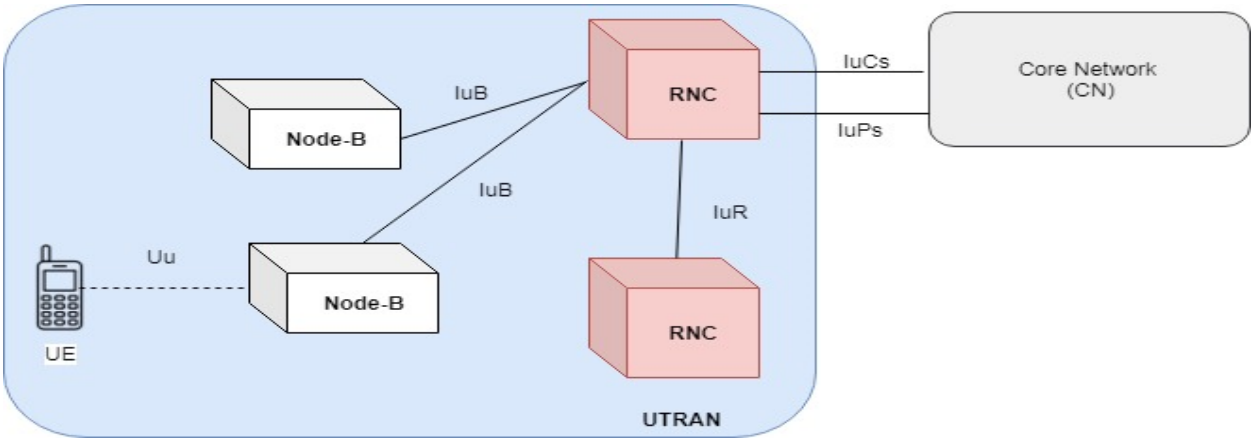

Figure 11. UTRAN overview with interfaces

Node-Bs are controlled by RNC and each UE is controlled by a particular RNC to further process the signaling to the core network (Bauer, Schefczik, Soellner, & Speltacker, 2003).

b) User Equipment

A 3G network terminal is the UE (User Equipment), as indicated in Figure 10. The UE consists of two separate entities, namely the Mobile Equipment and the UMTS Service Identity Module (USIM). In 2G GSM network it is known as Mobile stations (MS). UE is responsible for communication functions, which are on other side of radio interface. Following are the functions of UE for UMTS terminals (Bauer et al., 2003):

- Location update,

- Interface to an integrated circuit card for insertion of the USIM,
- Originating/receiving both connection-oriented and connectionless services,
- Support for the execution of algorithms for authentication and encryption.

USIM is also user dependent part of UE and is implemented into integrated circuit card UICC. UICC can have applications, software's or it can also have multiple USIM's. When the user subscribe to an operator in any case it will provide the information contents of USIM's (Bauer et al., 2003).

c) UMTS Core Network

UMTS core network (CN) is based on the GSM network architecture (Bauer et al., 2003), in which packet data services can be used effectively. This generates the profits and every operator has saved costs by utilizing this platform. GSM/GPRS network subsystems are capable of providing communication services for both circuit and packet switched traffic along with value added services (Eberspacher et al., 2001). Because of this, the 2G architecture has become the base for UMTS CN (see Figure 12).

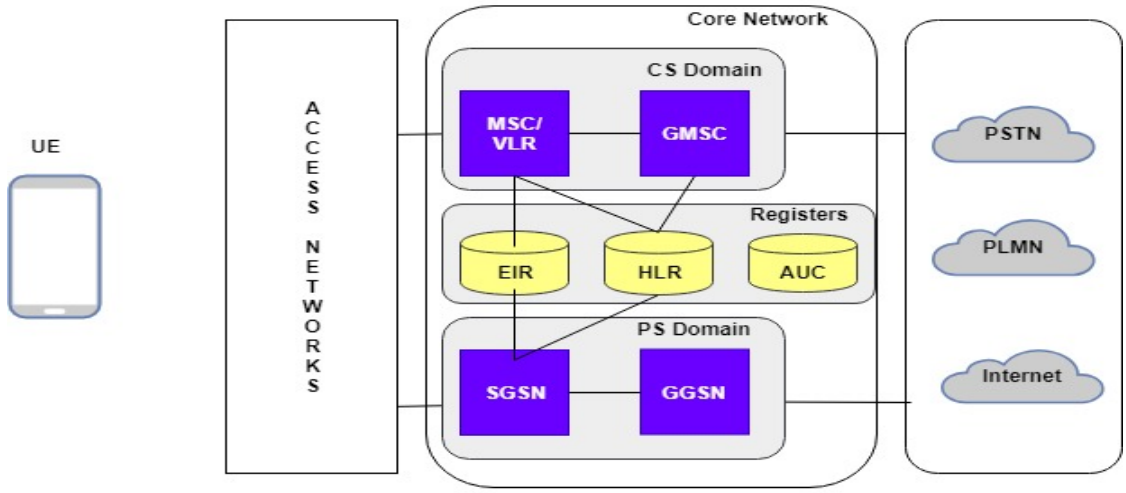

Figure 12. Core Network with domains

The Core Network is separated into two different domains: *Circuit Switch* domain (CS) and *Packet Switch* domain.

The Circuit Switch domain provides connection activities, mobility management, location update, location registration and security activities (Eberspacher et al., 2001). GSMC is part of the CS domain, which is responsible for incoming/outgoing connections to and from other networks. The difference between 2G and 3G in this domain is that the specifications of transcoders were part of radio network in 2G, but in 3G they are part of the core network (Bauer et al., 2003).

Packet Switch domain (PS) is also divided into two network elements, namely a Serving GPRS (SGPRS) and Gateway GPRS Support Node (GGSN). SGPRS communicates with the Access network over Iu interface in 3G, whereas in GSM it communicated over Gb (Eberspacher et al., 2001). This is related to the packet part of communication, which includes routing area update, location registration, controlling and security mechanisms etc. (Bauer et al., 2003; Kaaranen et al., 2005). GGSN is responsible for packet switching to other networks as shown in Figure 9 including the internet. HLR, EIR and AUC are considered in the Registers domain in CN of UMTS. It stores the information for both domains PS and CS of core networks (Kaaranen et al., 2005).

1.2.2. New Concepts in UMTS Network

a) Location services

A service is the essential unit that provides the information of exact MS. The location of UEs is an important factor, which provides the value added on top of the services to the user. For example, a list of hotels nearby after arriving in a particular city. In the case if someone is performing an emergency call, the place of ME should be known. There are three ways to determine the location, which are as follows (Korhonen, 2003):

- Cell-coverage-based method,
- Observed time difference of arrival (OTDOA),
- Network-assisted GPS methods.

b) High Speed Downlink Packet Access

High speed downlink Packet access HSDPA was designed for downlink data transmission in 3G networks. HSDPA was not part of release 3GPP 99, but was implemented in Release 5, which has data transfer rates up to 14.4 Mbps and upload 2 Mbps per cell (Korhonen, 2003).

c) Multimedia Broadcast/Multicast Service

The Multimedia Broadcast/Multicast service (MBMS) is a point-to-multipoint interface service for 3GPP cellular networks within the cell and core network. Mobile TV service is an example of MBMS. Transport bearers are used on which IP multicast packets are delivered (Hartung, Horn, Kampmann, Lohmar, & Huschke, 2009).

d) Gateway Location Register

The Gateway Location Register (GLR) is a location management service, which obtains the information of roaming subscribers in a visited network. Thus, in this case there is no involvement of HLR. UMTS is an international system that has reduced the PLMN signaling traffic due to GLR. This has saved the cost of long-distance international links. Every GLR serves one PLMN network (ETSI-TS 23.119, 2012).

1.2.3. Strengths and Weaknesses of 3G

Strengths

- Overcrowding is reduced with radio spectrum,
- Bandwidth is increased as compared to GSM. 144 Kbps is for rural, 384 kbps for urban outdoor and 2048 kbps is for indoor,
- It utilizes both WCDMA air interface and GSM infrastructure,
- Enhanced location-based services as discussed in section 1.2.2,
- Universal roaming.

Weaknesses

- High processing speed because of WCDMA, Power consumption is high,
- Although offering a higher data rate than 2G, it is still not sufficient,

- We are stuck while using the internet through mobile when inter-operability between different networks is obtained (D. S. Kumar & Meraj ud in Mir, 2015).

Table 2. Network structure of 2G and 3G

Feature	2G	3G
Core network	MSC/VLR, GMSC HLR/EIR/AUC	MSC/VLR, GMSC, HLR/EIR/AuC3G-SGSN & GGSN
Switching	circuit	Packet except circuit for air interface
Radio Access	(BTS, BSC, MS) FDMA, TDMA, CDMA	(Node-B, RNC, MS) W-CDMA, CDMA2000
Handsets	Voice based handsets	Multimedia terminals (voice, video, WAP)
Databases	HLR, VLR, EIR, AuC	Enhanced HLR, VLR, EIR, AuC
Data Rates	Up to 9.6 Kbps	Up to 2Mbps
Applications	Voice, SMS	Internet, multimedia

1.3. Long Term Evolution - LTE

3GPP started to define the objectives for LTE in 2004 in Toronto, which was set to satisfy the following requirements (Nohrborg, 2019):

a) Reduced cost per bit,

b) Simple architecture and open interfaces,

c) Flexibility usage of existing and future frequency bands,

d) Reasonable terminal power consumption,

e) Enhanced user experience-more services with lower cost and high speed.

When comparing to HSPA Release 6, the 3GPP LTE aims towards superior performance and targets set as below (A. Kumar et al., 2013):

a) 2 to 4 times more spectral efficiency than HSPA Release 6,

b) Peak rates beyond exceed 100 Mbps in downlink and 50 Mbps in uplink,

c) Round trip time < 10ms, to reduce latency,

d) Optimized packet switching,

e) High-level mobility and security,

f) Flexible frequency with 1.5 MHz to 20 MHz allocations.

1.3.1. LTE Technologies

The technologies utilized in 4G LTE are advances of the ones from the 3rd generation UMTS. The same reflect the efforts to achieve and realized the promised goals, while maintaining the best practices from before, that is, serving the end user with the best feasible Quality of Service (A. Kumar et al., 2013). Some of these technologies are OFDM (Orthogonal Frequency division Multiple Access) (A. Kumar et al., 2013),MIMO (Multiple input multiple output) (A. Kumar et al., 2013) and are described as under

Orthogonal Frequency Division Multiple Access (OFDMA) technology is used to achieve the larger bandwidth up to 20 MHz (A. Kumar et al., 2013). LTE defines several channel bandwidths, therefore the greater the channel greater is bandwidth. The access schemes are further divided into two approaches, which are being used in Downlink (DL) and Uplink (UL), respectively. For DL, 4G LTE uses Orthogonal Frequency Division Multiplexing (OFDM[6]) and Single Carrier - Frequency Division Multiplexing (SC-FDMA[7]) (Kussum & Malhotra, 2014). The overall motivation for OFDMA in LTE has been due to the following properties (Kussum & Malhotra, 2014):

- Good performance in frequency selective fading channels,
- Low complexity of base-band receiver,
- Good spectral properties and handling of multiple bandwidths,
- Compatibility with advanced receiver and antenna technologies.

MIMO (Multiple Input Multiple Output): Abbreviated as MIMO, is a wireless technology, which increases the RF radio's data capacity by using multiple transmitting and receiving antennas. It allows reliable and increase in data rate since data transmitted through multiple

[6] OFDM enables high data bandwidth transmission efficiently while resilience towards reflection and interference
[7] SC-FDMA technique used in LTE uplink transmission for better performance of Peak to average ratio (PARP) and frame error rate compared with OFDMA.

antennas over the path with same bandwidth. Due to multiple transmitters and receivers, it gives better signal strength even if the scenario is not in sight. There are less chances of signal loss and due to this feature quality of services at the user side is very high (S. Kanchi, 2013).

Figure 13 shows an implementation of MIMO that consists of transmit antennas and receiving antenna also known as 4x2 MIMO. MIMO can provide efficient performance with single antenna as compared with scattered radio waves and multipath between transmitting and receiving antennas. Different modes are standardized through MIMO is implemented such as single antenna, Transit diversity, spatial multiplexing, multi-user MIMO (MU-MIMO) (J. Lee, Han, & Jianzhong, 2009).

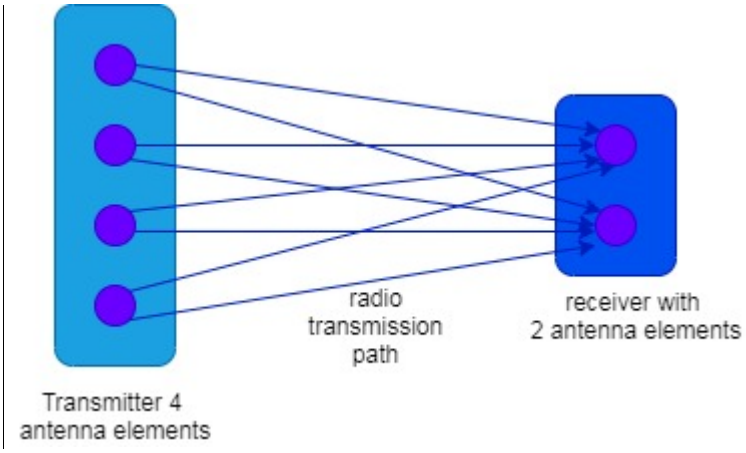

Figure 13. 4x2 MIMO (Brydon, 2013)

- Single antenna: Data is transmission is performed on one antenna and receiver can be single or multiple antennas, single antenna implementation mode is also known as SISO (single input single output) or SIMO (single input multiple output) (J. Lee et al., 2009),
- Transmit diversity: Another form of LTE mode that utilizes transmission of same data stream from multiple antennas. This LTE MIMO mode improves signal quality at receiver end and is used at channels modes i.e. control and broadcast channels (J. Lee et al., 2009),
- Spatial multiplexing: two or more antennas are used to transmit the information stream, and provide optimization of data stream by enabling the separation the data stream at receiver side (J. Lee et al., 2009),

- MU-MIMO: Performs spatial streams for different users in LTE MIMO. Spatial stream is multiplexing antennas of different spaces in a spectral channel, spaces are called spatial streams (J. Lee et al., 2009).

1.3.2. LTE Architecture

LTE architecture has evolved with a series of inventions from GSM to UMTS and resulted in LTE or the Evolved - Universal Terrestrial Access Network (E-UTRAN), introduced in (3GPP-TS 23.002, 2001). GSM works in a circuit switched environment with no support for IP. Data can be used in the later releases of 2G, but the data rate is very low. IP-based packet switched solution was introduced with an evolution from GSM to UMTS. UMTS (see Figure 12) consists of packet switching and circuit switching domains. When a data service established an IP address allocated to UE and released when the service is released, the issue was that incoming data services are still relying on the circuit switched core(see Figure 14), (Nohrborg, 2019).

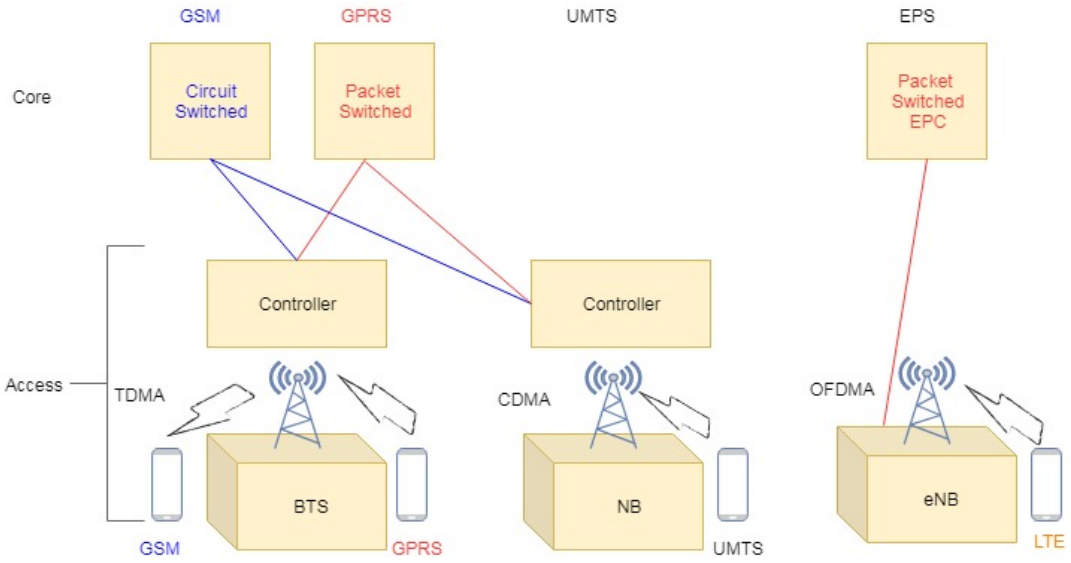

Figure 14. From GSM to LTE (Nohrborg, 2019)

Evolved Packet System (EPS) is a purely IP-based system, in which both voice and data services is carried by IP protocol. An IP address is allocated when the User Equipment is switched on and released when switched off. Figure 14 shows the evolution from GSM to LTE, whereas in Figure 15 the LTE architecture is depicted.

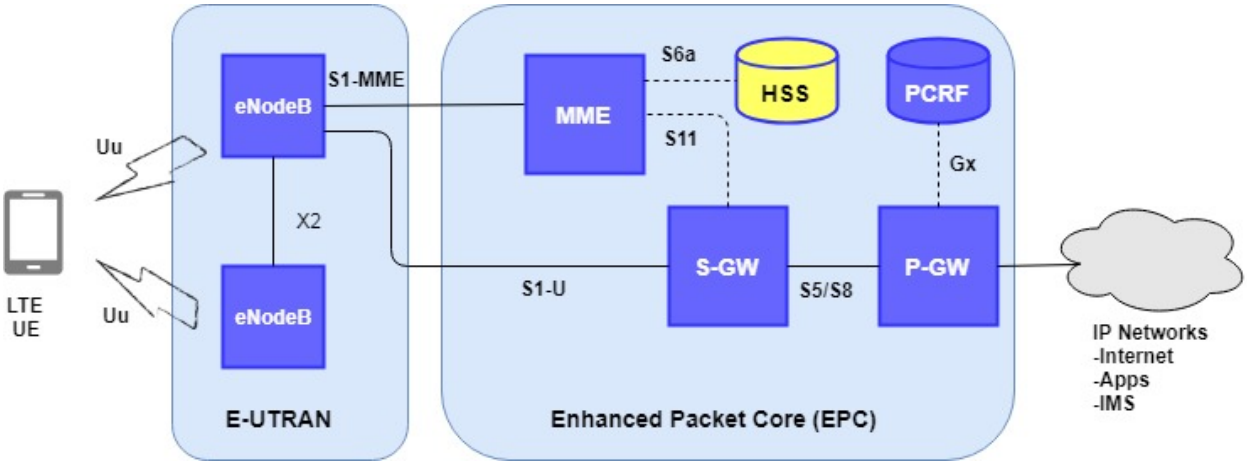

Figure 15. LTE architecture (Harri & Antti, 2009)

a) eNode-B E-UTRAN

The LTE Radio Access Network uses OFDMA in combination with higher modulation orders, up to 64QAM, as well as large bandwidths up to 20 MHz and MIMO spatial multiplexing in the DL (up to 4x4) for achieving higher data rates. The high data rate by utilizing MIMO can be as high as 300 Mbps in DL and 75 Mbps in UL. The LTE network is a network of base-stations called Evolved NodeB (eNode-B, or short eNB), with no centralized controller in between Enhanced Packet Core (EPC[8]) (Firmin & 3GPP, 2020) and NB as was the case with 3G UMTS. The eNBs are connected with X2-interface and S1 interface towards the core network (see Figure 16). Moreover, EPC separates user-plane and control plane, user-plane carries data traffic and control plane carries signaling traffic. TCP, UDP, IP packets are handled by user plane part while Radio resource control RRC protocol is part of control plane that transmit signaling message between base stations and the UE. By separating the user and signaling traffic enables flexible deployment of network in terms of scaling without affecting the live operations of existing nodes that are part of radio access network and EPC (Schmitt, Landais, & Yang, 2017).

[8] EPC is a key component of LTE architecture that consists of Serving gateway(S-GW), Packet data network (PDN), MME in order to separate user-plane and control plane data i.e. functional split (Firmin & 3GPP, 2020)

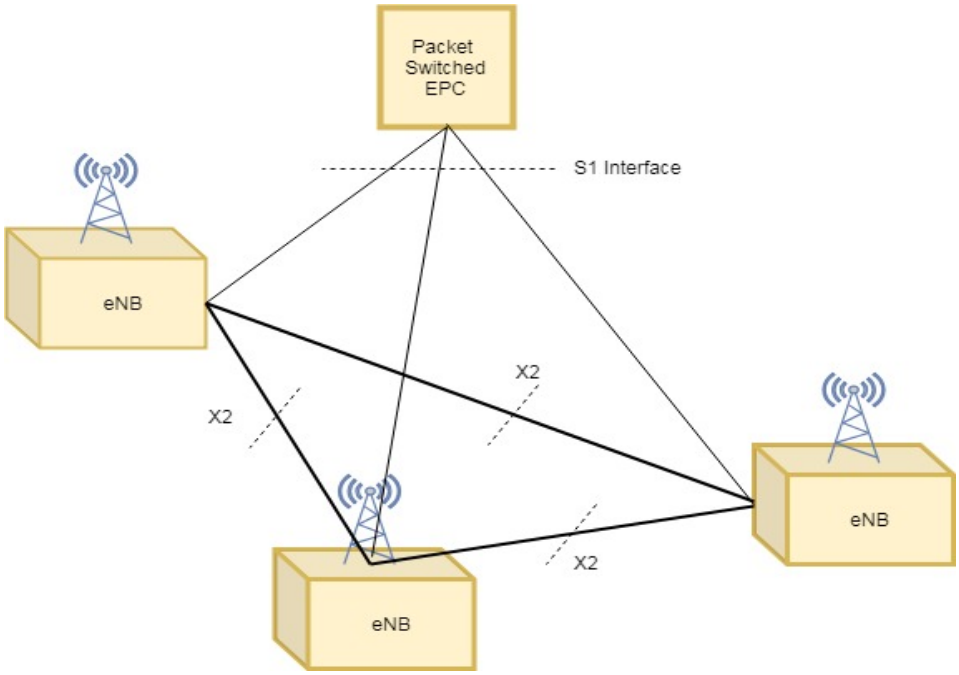

Figure 16. LTE network interfaces (Nohrborg, 2019)

LTE stresses an additional advantage at MAC protocol layer, which is used for scheduling between the eNB and UE and is part of user-plane protocol stack. Unlike that, in UMTS, the MAC protocol and scheduling is located at the controller (see Figure 14). The radio network part of LTE is referred to as E-UTRAN where 'e' is an abbreviation of evolved and has been added with NodeB as an eNB in LTE. There are three major elements of eNode-Bs as below (Nohrborg, 2019):

- The antennas, which are the most visible part of a mobile network,
- Radio modules that modulate and demodulate all signals transmitted or received on the air interface,
- Digital modules that process all transmitted and received signals on the air interface and that act as an interface to the EPC over a high-speed backhaul connection.

To reduce the cost of coaxial cables to the antennas, radio modules are installed near the antennas, also known as Remote Radio Heads (RRH) (Harri & Antti, 2009). Bearers[9] serve to commence a LTE radio network in a case where antennas and base stations cannot be installed in

[9] Bearers in LTE network are tunnels that provide connection between UE and Packet data networks (PDNs) through Packet network gateway (P-GW), by default UE has one bearer. (Juniper Networks, 2011)

each other's proximity (Juniper Networks, 2011). For that instance, a logical channel is made between network entities and described the QoS attributes, such as latency and maximum throughput data that flows on top. The overall transmission between mobile and radio base stations is managed via RAB (Radio access Bearer). This RAB further consists of Signaling Radio Bearer (SRB); once the connection is established with a mobile device, this SRB performs exchanging session management, mobility management and Radio Resource configuration (RRC) messages, and at least one Data Radio Bearer over which IP user-data packets are transferred. Correspondingly, LTE eNode-Bs are autonomous units, which do not require intelligent controllers as was the case with UMTS. We can therefore summarize that an eNode-B is not only responsible for providing air interface, but also manages the user management and scheduling of resources over the air interface, ensuring the QoS, mobility management and interference management. See Figure 17, which shows eNodeB connections to other logical nodes and main functions (Harri & Antti, 2009).

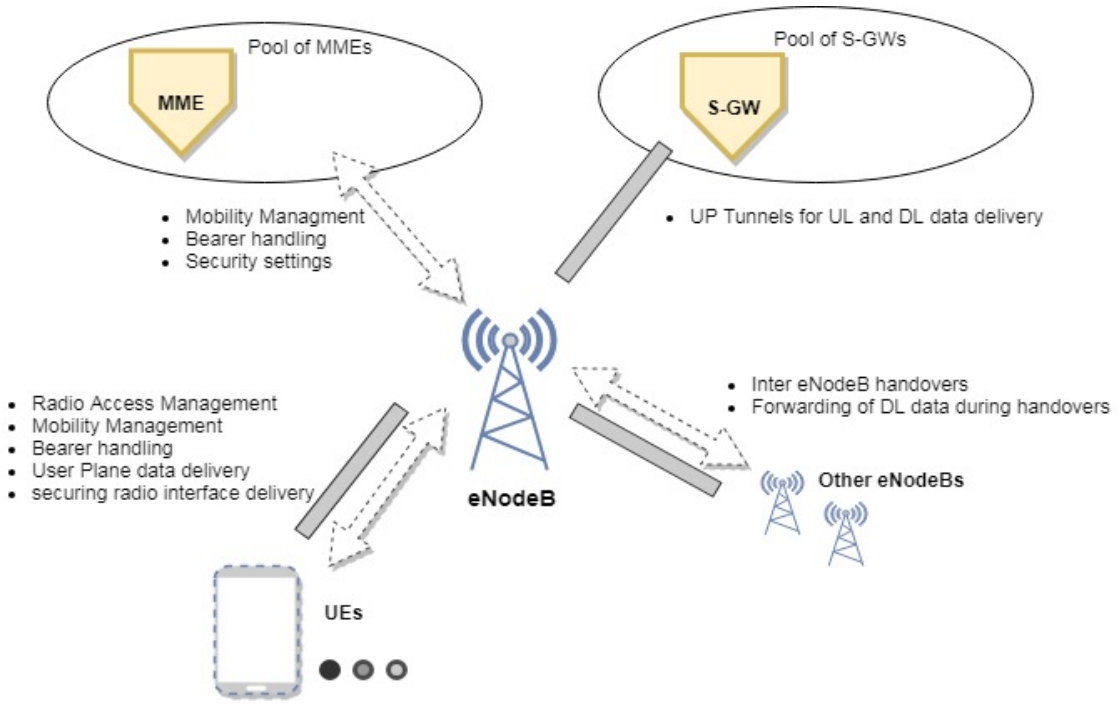

Figure 17 eNodeB connections to other logical nodes and main functions (Harri & Antti, 2009)

b) Mobility Management Entity (MME)

The Mobility Management Entity (MME) is the control element in the EPC core (Nohrborg, 2019). The main functions of MME in LTE core network are listed below:

- **Authentication and Security**: MME initiates the authentication by identifying the permanent identifier from the database or from UE itself. This function compares the parameters for the assurance that it is the same UE who is claiming to be. MME has ciphering and integrity protection algorithms, which compare the master key received in authentication vector from Home Subscriber Server (HSS). Diameter protocol is used in LTE EPC for exchange authentication, authorization and accounting (AAA) information which is more reliable and secure to legacy protocols such as (Remote Authentication Dial-In user service) RADIUS. HSS, PCRF functions use diameter protocol for information exchange (AAA) using IPsec or Transport layer security (TLS) protocol (Olsson, Mulligan, Sultana, Rommer, & Frid, 2012).

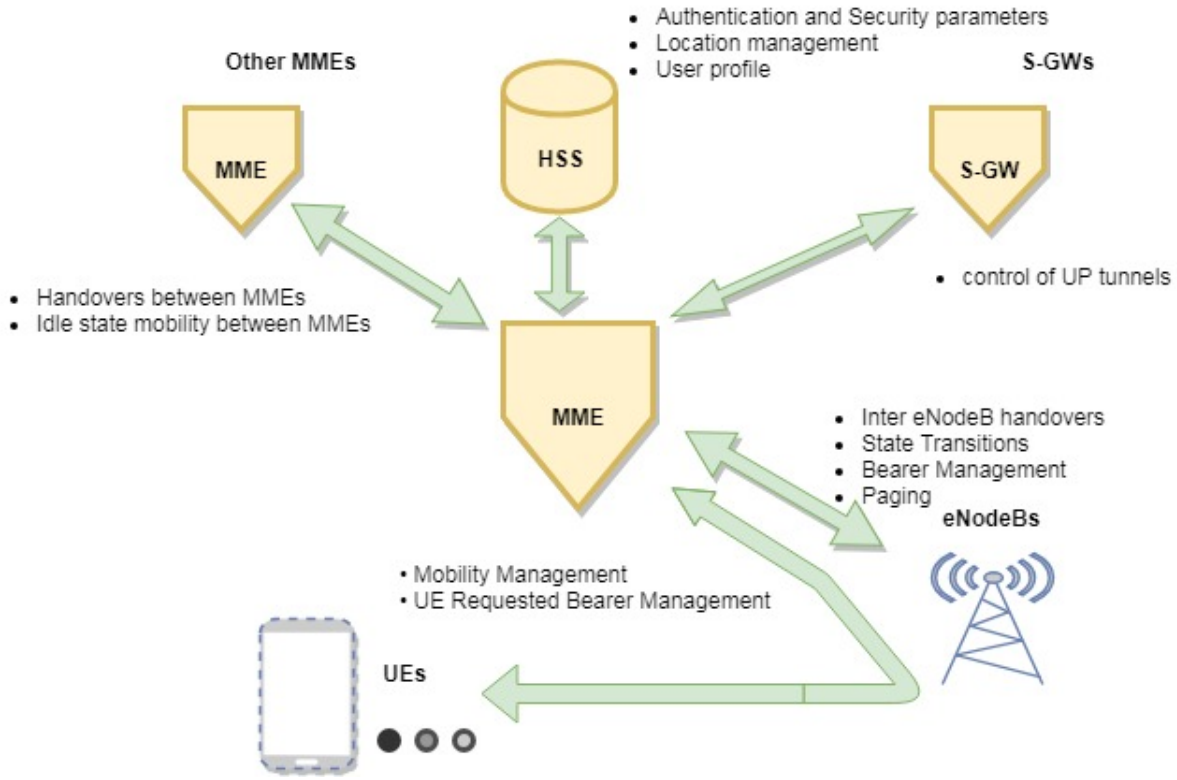

Figure 18. MME Functions (Harri & Antti, 2009)

- Mobility Management: MME also keeps track of the location of UEs in its service area (Harri & Antti, 2009). It will request appropriate resources to setup in eNodeB as well as S-GW, which it selects for UE. With this information, the tracking of active UEs is being performed by a Track Area (TA) group (defined in eNodeB). The resources are released once UE becomes inactive. Moreover, the MME also participates in control signaling for

handover of active UE between eNodeB, MMEs or S-GWs. Since there is no Radio Network Controller (RNC) in LTE network, the idle UE will update its location periodically, even when it moves to another TA. MME is being notified by external network for an idle UEs, where it requests available eNodeBs in the TA that is stored for the UE to be paged (Harri & Antti, 2009).

- Managing Subscription Profile and Service Connectivity: when UE is being registered to the network, the MME is responsible for fetching the subscription profile from the HSS. This profile will be stored in MME until the UE is inactive. When UE is active, this profile will ensure what data packet connection to be provided to UE at the network level. For setting the establishment of a dedicated bearer, MME receives request from either S-GW or operator service domain. Figure 18 shows the summary of MME connections and main functions (Harri & Antti, 2009).

c) Serving Gateway (S-GW)

The S-GW is the termination point of the packet data interface towards E-UTRAN (Harri & Antti, 2009). When terminals move across eNodeB in E-UTRAN, the Serving Gateway acts as a local mobility anchor, meaning that packets are routed within intra E-UTRAN and mobility is established with other 3GPP technologies, such as 2G/GSM and 3G/UMTS (Harri & Antti, 2009).

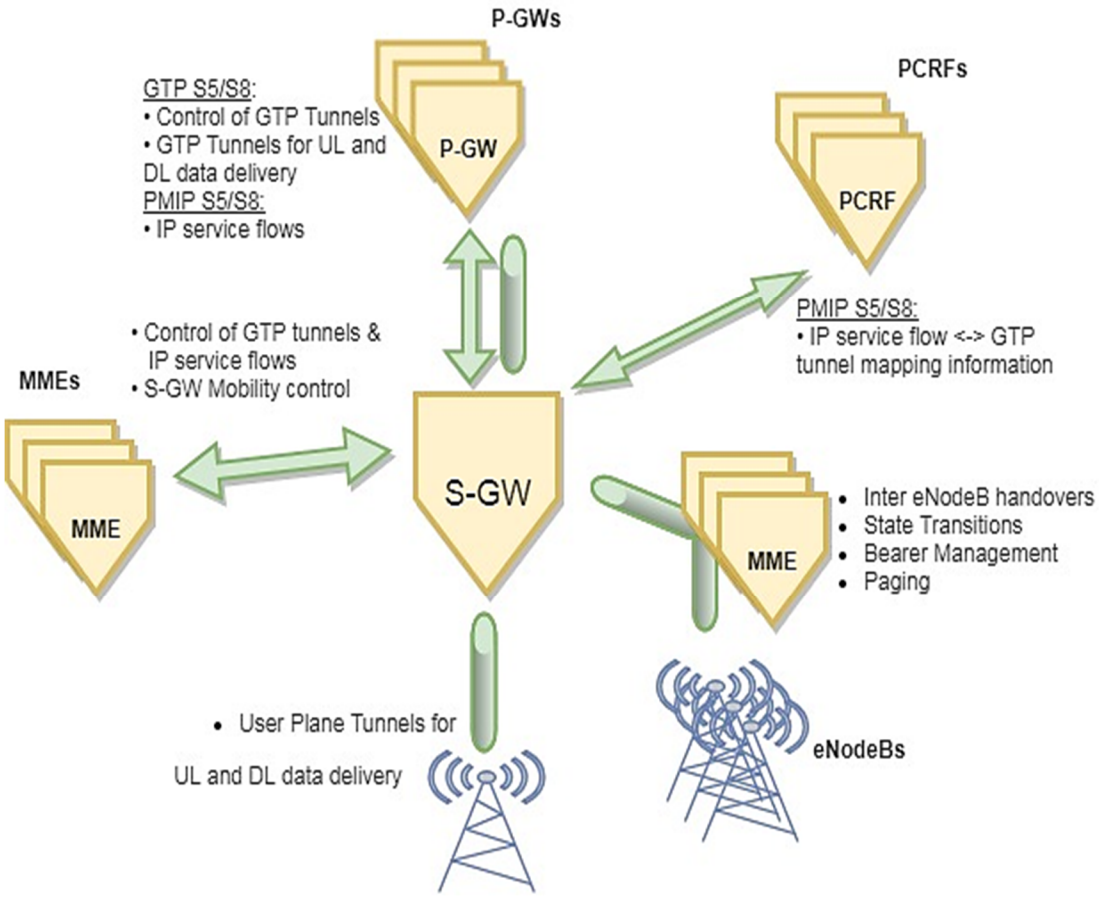

Figure 19. S-GW connections and main functions summary (Harri & Antti, 2009)

Another function of S-GW is tunneling and mapping the IP service flow. The S5/S8 interfaces are set between S-GW and P-GW, which have GTP[10] tunnels for User Plane (UP) (Olsson et al., 2012). Mapping between IP service flows and GTP tunnels is done in P-GW with GTP on S5/S8 interfaces. The S-GW performs mapping between IP service flows in S5/S8 and GTP tunnels in S1-U interfaces. The mapping information from Policy and Charging Resource function (PCRF) (described on page 39) it receives when (Proxy mobile IP protocol) PMIP[11] is utilized by S5/S8 interfaces. Figure 19 shows the S-GW connections to other logical nodes and main functions. One S-GW may serve one area and limited number of eNodeBs. When UE moves, the S-GW can be changed but will have the same P-GW in a whole network, as P-GW does not change during mobility. In case of Indirect forwarding, there is no specific interface since the format is same as in S1-U interface, considering all S-GWs are communicating directly with eNodeB.

[10] GTP: GPRS tunneling protocol used for encapsulation of packet between core elements. In user place (GTP-U) will encapsulate the UE TCP/IP packet and forward to S-GW,

[11] PMIP(mobility protocol) is ip interface(s5) between S-GW and P-GW supports non-trusted 3GPP network. (Olsson et al., 2012)

d) Packet Data Network Gateway (P-GW)

Listed below are the functions of the PGW (Packet Data Network Gateway):

- It provides connection between UEs and Packet data Network (PDNs),
- UEs can connect to multiple PDNs with multiple PGWs but will still be served by one S-GW (see section 1.3.2 c),
- P-GW performs the required DHCP functionality or queries an external DHCP server, which delivers the address to UE,
- It also performs gating and filtering functions as per policies set for UE (Harri & Antti, 2009).

e) Policy and Charging Resource Function (PCRF)

Policy and Charging Resource Function (PCRF) is an element of LTE network architecture in the Evolved Packet Core. It is responsible for Policy and Charging Control (PCC) which decides how to handle the services according to QoS and pass information to PCEF located P-GW. This leads to setup bearers and policies accordingly. These bearers are set up when the UE initially connects to the network.

f) Home Subscriber Server (HSS)

Home subscription server (HSS) is a database for all the permanent users. It has all information related to the user of visited network control node, for example MME. In addition, HSS stores the master copy of subscriber profile services and allowed PDNs connections. Furthermore, roaming and visited networks can be allowed or disallowed and stored in HSS. Authentication Center (AuC) is a part of HSS in a LTE network, which has a permanent key to calculate the authentication vectors that are exchanged with the visited network. This serves the purpose of authenticating the UE and allowing them to move migrate from one MME to another. When UE is active, HSS will point to that serving MME at a particular time and when UE is moved to another MME, HSS will cancel the location of the previous MME (Harri & Antti, 2009).

1.3.3. Strengths and Weaknesses of LTE

Strengths

- LTE provided, with the major improvements, peak bit rates within 64QAM and 2x2 MIMO, from 14Mbps to 42Mbps,
- UE power consumption is reduced due to discontinuous transmission and reception,
- Has reduced the network latency and has improved the IP based services,
- Capability of providing speeds from 100Mbps to 1Gbps,
- Supports broadband access, mobile TV, HDTV content.

Weaknesses

- 4G is hard to implement and requires complacent proprietary and expensive hardware,
- Although security has improved in 4G, the overall architecture is flatter and less hierarchical, without RNC. Thereby, it is more exposed to DDOS attacks,
- The usage of higher frequencies, while providing higher data rates, results to smaller cell sizes, hence a higher number of base stations,
- There is a lot of burden during interworking scenarios with 2G/3G networks due to handovers, cell changing and signaling.

Table 3. Summary 4G/LTE

Deployment	2000-2010
Bandwidth	200Mbps
Technology	IPv4&6:LAN/WAN/WLAN/PAN
Service	High definition streaming, Probability increased
Multiplexing	DL:OFDMA , UL:SC-FDMA
Switching	All packet
Core network	Internet
Handoff	Horizontal and vertical

1.4. Fifth generation networks – 5G

The fifth generation 5G is the next evolution step after LTE (5GPPP, 2019). A tremendous change was evident from the previous legacy 1G to 2G networks and LTE. Accordingly, 5G is expected to bring significant change in the modern world through advanced digital transformation and being able to connect various devices without any constraints in an efficient and smart way.

1.4.1. 5G expectations and key drivers:

The 5GPP research program is building the milestone of future networks and use-cases (i.e. eMBB, mMTC, URLLC) from industrial sector to transform the business models into digitalized and virtualized (5GPPP, 2019). In the year 2015, the European commission agreed on key challenges to be achieve the fifth generations of network (5GPPP, 2019).

- **The wireless capacity** to be increased 1000 times as compared with 10 times in 2010 by introducing new frequencies in the mmWave spectrum. It will be a ubiquitous network where every person, device/machine will be connected to 7 billion people (5GPPP, 2019).
- **Saving energy to 90%** although the wireless capacity shall increase, it represents a significant challenge to reduce energy consumption where 7 billion people will access the service ubiquitously (5GPPP, 2019).
- **7 Trillion Connecting things** goal includes internet of things (IoTs) and millions of sensors that will communicate 7 billion people or other machines in a network and continuously using specific services as per user/industry requirements (5GPPP, 2019).
- **Zero Latency** in providing the services to 7 billion people and 7 trillion of things for smooth running of operation and provide home-like environment to users to access services with high speeds (5GPPP, 2019).

To achieve above expectations three key use cases are defined by 5GPP (5GPPP, 2019), 3GPP (3GPP-TS 23.501, 2018), GSMA (Gsma, 2019a) that will drive to standardize 5G. These are:

- Enhanced Mobile Broadband
- Critical Machine-type communication

- Massive machine to machine communication

Enhanced Mobile Broadband (eMBB)

Enhanced Mobile Broadband will provide high data-rate speed with low latency communication (NGMN, 2015). It will involve wide area coverage which will be 1000 times more compared with legacy network LTE i.e. 10 times in 2010. Urban areas, rural areas, offices, indoor/Outdoor customers would access eMBB from anywhere. It is expected the downlink speed could reach up to 50Mbps when outdoor and 1Gbps for indoor environment. Similarly, uplink is expected to have 25Mbps in outdoor and 50 Mbps in indoor environment, respectively. Ultra-high bandwidth will be a challenge of eMBB since smart offices will be connected wirelessly (NGMN, 2015), Infrastructure having ultra HDV conferences and group meetings and applications interaction is a challenge to cope in 5G network. Operator cloud services, HD videos sharing from concert or stadium are some of the other which require ultra-high connection density, high data rate and low latency. Moreover beyond 2020 there will be an increase demand for mobile services in airplanes, bullet trains and in vehicles. Nowadays the high speed trains travels with 500km/passengers will use UHDV, online gaming services, accessing cloud and even accessing company systems and performing real-time applications development while traveling with such high speeds (NGMN, 2015). It will be a challenge to provide such environment with eMBB. In this use case family capacity variation will also be a big challenge to cope with stationary to busty traffic.5G is expected to cope with stationary, non-stationary, dynamic, and real-time provisioning of capacity radio network.

Ultra-Reliable Low Latency Communication (URLLC)

URLLC is the second use case of group that is commonly agreed by all industrial standards (5GPPP, 2019). The vision of 2025 suggests there will be an immense growth in automotive, health, transport sections (Gsma, 2019a), with prediction that the manufacturing and agriculture sector will fully rely on real time systems communication, This may involve completing job remotely which require zero to very low latency. Automated traffic control and driving will be fully implementable only if 5G provide reliability, zero to low latency in this scenario. This application will facilitate the communication of vehicle to vehicle and road users for safety and smooth flow of traffic. Automation in a control network of robotics is another use case family in URLLC in which they are used for manufacturing of services with a reaction time of 1ms or less.

Remote health monitoring and treating patients is also one of the vision of 2020+ (NGMN, 2015). There will be several real time health applications with sensors for the surveillance of patient entities remotely. Surgical operations through advanced robotic technology will be implemented remotely soon. 5G standards must assure the authentication, privacy, low latency for each device communication.

Drones is another domain that will be used for logistics purposes which include autonomous logistics in densely or rural areas. Drone taxis is another domain which require the device to device communication in a 3D connectivity behavior from air to basements of buildings and subways taking advantage of 5G communication. In a nutshell enhanced machine to machine communication URLLC requires 5G to be reliable, secure, and scalable and has a feature of mobility and capacity to facilitate such an environment (NGMN, 2015).

Massive Machine to Machine Communication (mMTC) and Internet of Things (IoT)

The vision of 2020 also includes the third use case of massive communication of devices which is depicted 7 trillion by 8 billion people defined by (5GPPP, 2019) as mMTC (massive machine to machine communication). This use case deals with the data and demands of sensors, actuators and cameras and human-type communication (HTC). Internet of things (IoTs) are taken into consideration that will exploit the 5G framework.

Smart Wearable is an example, which consists of multiple sensors to measure human blood pressure, body temperature, heart rate etc. the challenge is the management of such sensors that will be used by billions of people and the actual application associated with these IoT devices. Sensor networks integration is yet another example, which includes gas, electricity, and water metering devices, as the communication on this layer will be a challenge for a 5G internetworking framework. These applications and sensor-related machine-to-machine and machine-to-human communication, require highly reliable and secure network and mobility to be supported in the underlying 5G infrastructure (NGMN, 2015).

1.4.2. 5G Next-Generation Core Architecture

The 5G system consists of User Equipment (UE), Access Network (next-generation Node-B, or gNB) and Next-Generation Core Network (5GC). 5G exploits well-established techniques of Software-Defined Networking (SDN), along with open application programming interfaces

(APIs) to provide virtual Network Functions (vNFs), tailored and customized according to a specific scenario. Based on the use-cases, the architecture of 5G is represented in a substantially different demeanor, but, it is an evolution of the 4G system. Figure 20 represents an overview of the 5G architecture, which is separated into two main group functions: Control Plane (CP) core of 5GS and User Plane (UP) functions. The control-plane group has different elements defined in terms of Network Functions (NF). It is comprised of a common framework and offers services to other authorized NFs or users. This approach allows modularity, scalability, and reusability. Due to the nature of services in terms of functions, the 5G core architecture is also known as Service-Based Architecture (SBA) (Khan, 2018).

In a service-based architectural approach, application components provide services using communication protocols over a network. This makes the infrastructure independent of vendors, products and technologies. Moreover, the service functions can be accessed remotely, updated independently without any additional hindrance.

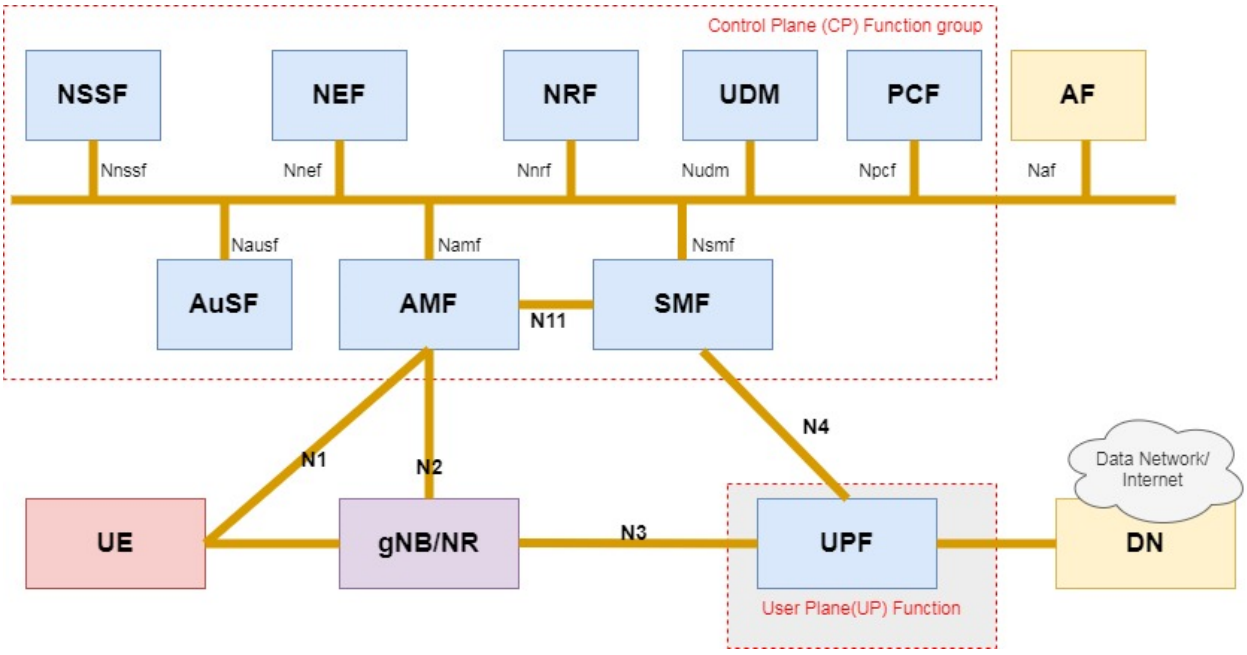

Figure 20. 5G core architecture

The 4G EPC core architecture (Figure 15) included the eNB, SGW, P-GW and MME modules. Each of these nodes are having proprietary hardware and software in a device, and there are too many various protocols into action. For every new system, a new signaling protocol is required between two interfaces. For example, in case of MME and VLR, SGsAP is the signaling

protocol that is being used only between them. This complexity of nodes and protocols leads to signaling storms (Khan, 2018).

Among the problems within the EPC, there was lack of gateway selection flexibility, as well as closed APIs. Therefore, in the big picture, 5G architecture abandons most wireless signaling protocols, replacing them with (Hypertext transfer protocol v2) HTTP2[12] (IETF, 2015) transport and(JavaScript object notation) JSON[13]. As discussed above, changing APIs framework from close to open, will allow read/write requests to any of the network functions independently. Additionally, we describe the network functions of the 5GC from Figure 20 (3GPP-TS 23.501, 2019).

a) Network Functions Functionality

i. User Plane function (UPF)

From Figure 20, we can observe the 5G UE connected to gNB or the Next Generation Radio Access Network (NG-RAN). The gNodeB connects to the User-Plane Function (UPF) via N3 interface. This function provides the necessary user-plane processing in the 5GC, the same way as it is done in the case of S-GWU (Serving Gateway User-Plane) in CUPS[14] architecture of the EPC (3GPP-TS 23.501, 2019). In addition, it has the functionality of PGW-U respectively. UPF also connects with Data network (DN) via N6 interface, which is similar to Gis interface in EPC. The third interface N4, is connected to the Session Management Function (SMF). Further functionalities of UPF are as follows (3GPP-TS 23.501, 2019):

- Packet routing and forwarding,
- Inspecting packets and lawful interception,
- QoS handling for User Plane (e.g. UL, DL rate enforcement),
- Policy enforcement of UP (packet filtering, gating, traffic forwarding),
- Responding to ARP requests by providing MAC addresses corresponding to IP address request,

[12] HTTP2 network protocol to define the framework of transmission of data over internet.(https://tools.ietf.org/rfc/rfc7540.txt),

[13] JSON framework for data-interchange between servers that is in human readable format. (Stackoverflow, 2020)

[14] CUPs provides separation of EPC(SGW & PGW) elements into user plane and control plane functions of gateways in result enables flexible network deployment, (Schmitt et al., 2017)

- Act as an external PDU session point of interconnection between data network (DN) and anchor intra/inter Radio access terminal (RAT) mobility (3GPP-TS 23.501, 2019).

In short summary, the 5G User plane function (UPF) is a S-GW-U + P-GW-U of the 4G architecture functionalities.

ii. Session Management function (SMF)

This function is in the control plane of the 5G core, which defines the architecture as service-based architecture (SBA). SMF is reorganized in 5GC by dividing the MME, SGW-C and PGW-C from EPC(3GPP-TS 23.501, 2019). Some functionalities of said entities in EPC are moved to SMF and the Access & Mobility function (AMF). In addition, the Session Management functionality of MME is included in SMF. Other functionalities are (3GPP-TS 23.501, 2019):

- IP address allocation to UEs and management,
- Session management (NAS signaling session, establishment of a session, modifying and release),
- DHCP services,
- Controlling UPF to proxy ARP and IPv6 solicitation requests by giving MAC address to IP sent in request,
- Gives path to traffic routing to proper destination at UPF,
- Downlink data notification,
- Support for communication related to user plane services, by determining the policy of charging data and interfaces (3GPP-TS 23.501, 2019).

In other words, the 5G SMF function is a complementary MME (Session management functionality) together with the control-plane functions of SGW-C and PGW-C.

iii. Access and Mobility Management function (AMF)

This function connects with UE via N1 interface and gNodeB with N2 interface. All UE related functions are managed by AMF. EPC functionalities of MME, SGW-C and PGW-C

(access & mobility) are done by AMF (3GPP-TS 23.501, 2019). Further functionalities of AMF are:

- Mobility and control management,
- SMF selection between UPF and UE,
- Access authentication and authorization,
- Interaction with the SMSF service to transport SMS messages,
- Mobility event notification,
- Support specific network slice authentication and authorization,
- Ciphering and integrity of NAS,
- Location service management for regulatory services (3GPP-TS 23.501, 2019).

It is important to note that there can be a single or multiple AMF functions per UE, but there is only one NAS interface instance per access network between the UE and the core network, terminated at just one of the AMFs that implements the NAS integrity and mobility management.

Figure 21 shows the summary of some main NF of 5GC and the evolution from 4G EPC (3GPP-TS 23.501, 2019).

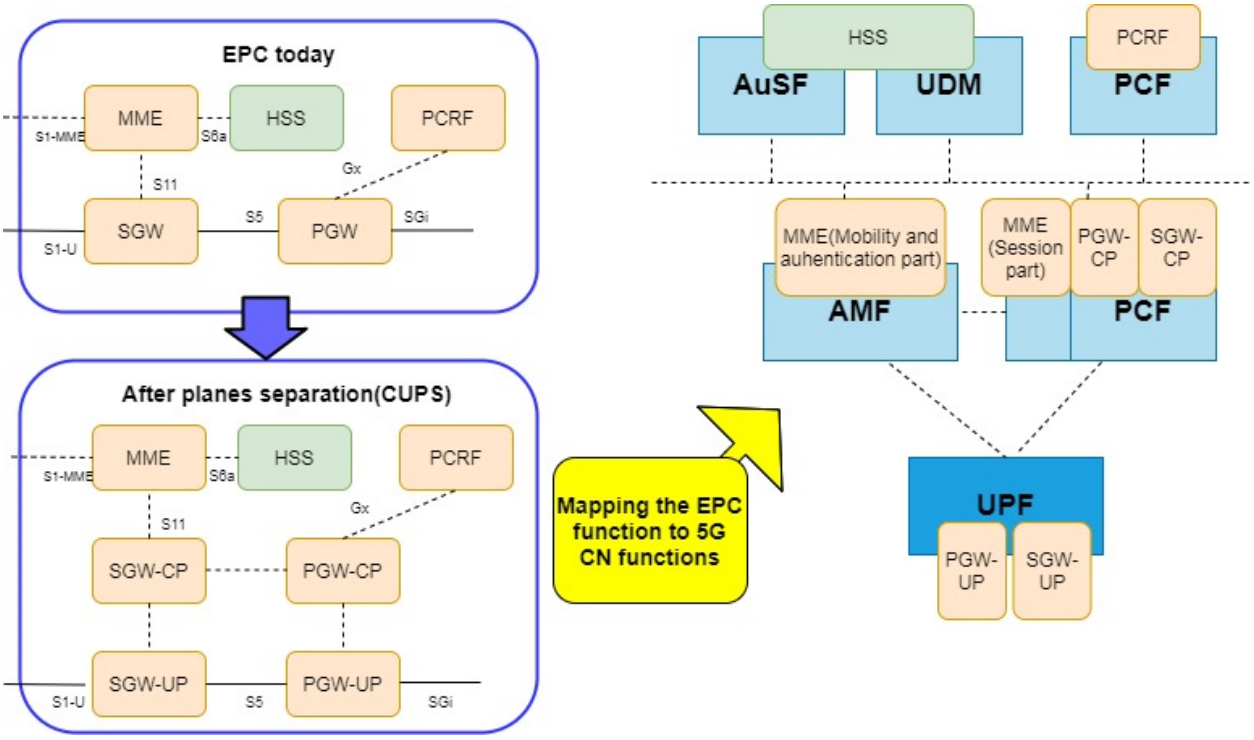

Figure 21. Evolution from 4G core to 5G core functions (Medium, 2018)

iv. Network Slice Selection Function (NSSF)

The Network Slice Selection Function (NSSF) supports the following functionality (3GPP-TS 23.501, 2019):

- Selecting the set of Network Slice instances serving a UE,
- Determining the Allowed NSSAI and, if needed, the mapping to the Subscribed S-NSSAIs,
- Determining the Configured NSSAI and, if needed, the mapping to the Subscribed S-NSSAIs,
- Determining the AMF that is set to serve the UE, or, based on configuration, a list of candidate AMF(s), possibly by querying the NRF (3GPP-TS 23.501, 2019; NGMN, 2019).

As indicated previously, a network slice is a logical network consisting of all required resources that are dedicated or shared, physical, or virtual on top of the same or different infrastructure. Slices can be independent, but they may also share resources. Computing, Access

equipment, transport, VNFs, processes and storage are some examples of resources. These slices can be employed with a strong notion of security, isolation of functions, QoS, policy enforcement, reliability, integrity and so on. 3GPPP has defined two types of identifiers in slicing scenario (3GPP-TS 23.501, 2019). One is S-NSSAI, which describes Single Network Slice Assistant Information; it is a kind of identification of slice, which gives information of a service type. S-NSSAI contains two parts: SST (service slice type) and SD (slice differentiator).

$$\text{S-NSSAI} = \text{STT} + [\text{SD}]$$

SST has been assigned 4 values that show the behavior in terms of features and services (Table 4).

Table 4. Standardized SST values

SST value	Slice/Service Type
1	eMBB
2	mMTC or URLLC
3	mIoT
4	V2x (3GPPP Rel-16)

A SD is an optional description to differentiate among multiple slices of the same slice/service type (3GPP-TS 23.501, 2019).

v. Network Exposure Function (NEF)

This function provides a framework for securely exposing services and features of the 5GC. With this exposure, NEF allows developers to compute new services that allow end-to-end services using exposed services and features by 5GC. External AF uses this framework through NEF. During the internal and external translation of functions, NEF stores information to Unified Data Repository (UDR[15]), which is used by other NF and AF for the purpose of analytics (3GPP-TS 23.501, 2019).

[15] UDR comprises of subscriber information that can be accessed by other functions such as PCF (policy control function) and Unified Data Management (UDM) function. (3GPP-TS 23.501, 2019)

vi. Network Repository Function (NRF)

NRF provides discovery function services to all other network functions. NF in 5GC can use the services of another NF directly, without passing any another node. Figure 22 illustrates a NRF service registration and discovery workflow.

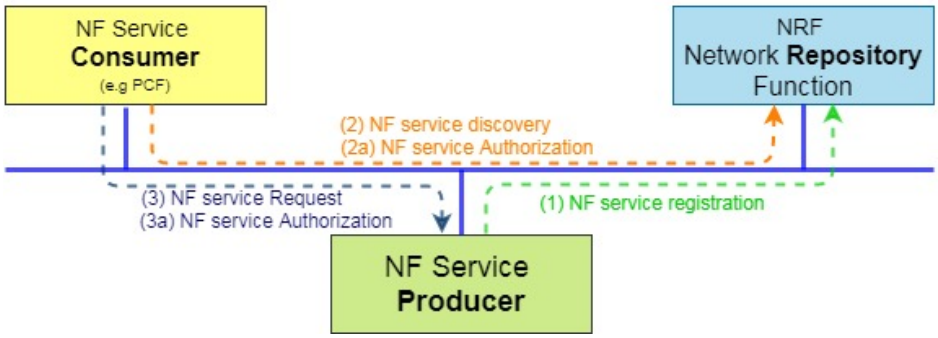

Figure 22. NRF service registration and discovery (Medium, 2018)

vii. Authentication Server Function (AuSF)

AuSF acts as an authentication function, which implements EAP authentication server. It also stores the public and private keys (3GPP-TS 23.501, 2019). Unified Data management (UDM) described later host subscription management and access authorization functions i.e. Authentication Credential Repository and Processing Function (ARPF) which allows interconnection with AuSF to exchange information (authentication data, keying) if required.

viii. Unified Data Management (UDM)

The UDM function is equivalent to HSS functionalities in EPC architecture.

- It includes support for generating keys and authenticating credentials,
- User identification and handling,
- Subscription management and access authorization (3GPP-TS 23.501, 2019).

1.4.3. Next Generation Radio Access Network Architecture

The 5G Next-Generation Radio Access Network (NG-RAN) consists of gNBs connected to 5GC architecture (see Figure 20), through NG interface. Figure 22 is an illustration of 5G-RAN architecture of the described functions (3GPP-TS 33.102, 2012):

- Radio Resource Management Functions: Radio Bearer Control, Radio Admission Control, Connection Mobility Control, Dynamic resource allocation to UEs in both uplink and downlink (scheduling);-IP header compression, data encryption and data integrity protection,
- Selection of an AMF at UE attachment when no routing to an AMF can be determined from the information provided by the UE,
- Routing of User Plane data towards UPF(s),
- Routing of Control Plane information towards AMF,
- Connection setup and release,
- Scheduling and transmission of paging messages,
- Scheduling and transmission of system broadcast information from AMF,
- Measurement and measurement reporting configuration for mobility and scheduling,
- Transport level packet marking in the uplink,
- Session Management,
- Support of Network Slicing,
- QoS Flow management and mapping to data radio bearers,
- Support of UEs in RRC_INACTIVE state,
- Distribution function for NAS messages,
- Radio access network sharing (3GPP-TS 33.102, 2012).

a) **The gNodeB (gNB)** is a logical next-generation radio access node. It is further separated in user-plane and control-plane protocol stacks that terminate to UE, as well as NG interface to 5GC. gNBs are interconnected with the Xn interface. As shown in

Figure 23, NG is further logically categorized in NG user-plane interface (NG-U), connected to UPF (User Plane Function) and the NG control-plane (NG-C) to the Access and Mobility Management Function (AMF) of 5GC (3GPP-TS 38.401, 2020).

b) **The gNB-CU or gNB control unit** is a logical node of gNB that operates single or multiple gNB-distributed units (gNB-DU). It has hosting protocols Radio resource control (RRC), SDAP and PDCP protocols to host gNB-DU via F1 interface, connected to gNB-DU for the purpose of optimization and performance RAN functions. gNB-CU is further divided into CP and UP part, as shown in Figure 23 (3GPP-TS 37.470, 2020).

c) **The gNB-DU or gNB distributed unit** is a logical node that supports one or multiple cells. One cell is connected to one gNB-DU and further it connects with gNB-CU via F1 interface. It has a protocol stack of RLC, MAC and PHY layers of gNB. gNB-DU operation is partially controlled by gNB-DU. The task of this node is to broadcast the system information by performing the encoding of NR-MIB and SIB1 messages (3GPP-TS 37.470, 2020).

d) **The gNB-CU-Control Plane (gNB-CU-CP)** is another logical node with RRC and PDCP (control-plane segment) of gNB-CU for gNB. E1 interface is terminated with gNB-CU-User Plane (gNB-CU-UP) and F1-C to gNB-DU.

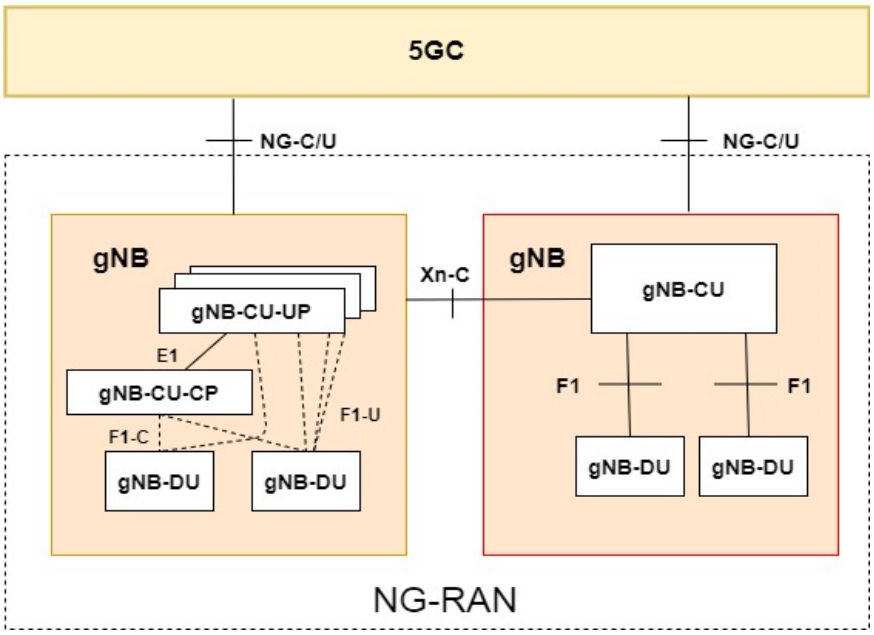

Figure 23. 5G NG-RAN architecture (courtesy 3GPP): (3GPP-TS 38.401, 2020)

e) **A gNB-CU-User Plane (gNB-CU-UP)** is a logical node hosting PDCP (user plane part), SDAP (user plane part) protocol of gNB-CU for gNB. It is connected to the E1 interface to gNB-CU-CP and F1-U connected with gNB-DU.

f) **The NG interface** is also divided into two types i.e. User-plane protocol NG-U, which has PDU session service and carries user data through the AS (Access stratum). The second type is NG-C, used for controlling PDU sessions and connection between UE and the network.

g) **F1-C (Control plane) Function** has four parts of functionalities. The first part is F1 Interface Management Function that performs gNB-CU configuration update, gNB-DU configuration update, indication of error and reset function. The second part is System Information and Management Functions, which provides signaling support for on-demand SI delivery and enabling UE saving energy. The third functionality is UE context[16] Management Function that is responsible for UE context acceptance, rejection, or modification. It also includes the establishment, modification, and release of Data Radio bearers (DRBs) and signaling Radio Bearers (SRBs). Last but not least, the RRC message transfer function is part of this interface, which is responsible for exchanging RRC messages between gNB-CU and gNB-DU (Bertenyi, Burbidge, Masini, Sirotkin, & Gao, 2018).

h) **F1-U (user plane) Function** allows user data to be transferred between gNB-DU and gNB-CU. Second functionality is Flow control, which allows us to control the downlink user data transmission towards gNB-DU.

i) **E1 interface** has E1 interface management and E1-bearer context management function.

It is also important to note that a gNB can have a single gNB-CU-CP, multiple gNB-CU-UPs and multiple gNB-DUs. A set of gNB-DUs can be connected to multiple gNB-CU-UPs with the main control node of gNB-CU-CP. And one gNB-CU-UP can be connected to multiple gNB-DUs under the control of the same gNB-CU-CP (Bertenyi et al., 2018).

[16] UE context is a set of information that requires UE to be in active state and logical connections S1 to maitain services between gNB and UE after completion of handover (3GPP-TS 38.473, 2018).

1.4.4. New-Radio Interface Protocol

This section discusses the protocols that operate between NG-RAN and the UE in the radio interface. Protocol architecture has a user-plane used for user data between UE and the network, as well as a control-plane that provides signal control between UE and NG-RAN. Figure 24 shows the protocol stack of user plane terminated in gNB on the network side (Bertenyi et al., 2018).

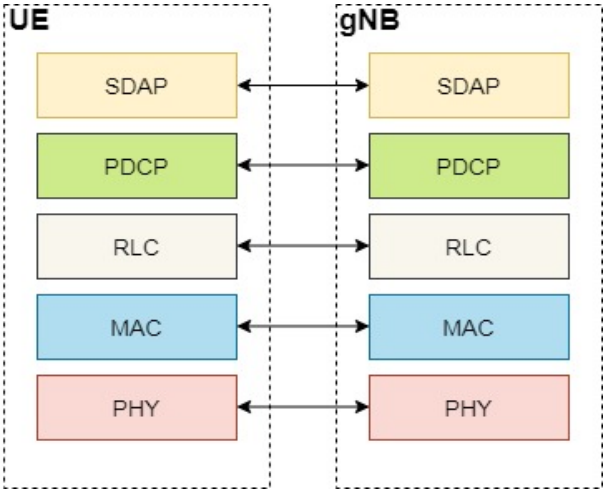

Figure 24. User Plane Protocol stack (Bertenyi et al., 2018)

a) User plane

i. Service data adaptation protocol (SDAP)

With SDAP protocol, the core network can configure QoS of IP flows in a PDU session[17]. This protocol provides mapping of IP flows from different QoS requirements and radio bearers to deliver the appropriate QoS (Bertenyi et al., 2018).

ii. Packet Data convergence protocol (PDCP)

PDCP protocol provides header compression and decompression through "Robust Header Compression" (RoHC), ciphering/deciphering and integrity protection. Furthermore, it also performs duplication detection and reordering of both transmitted and received PDCP PDUs. In the New Radio (NR) for 5G, data duplication over multiple transmission paths has been

[17] PDU (Protocol data unit) connectivity between UE and data network such as IPv4, IPv6 data (Bertenyi et al., 2018).

introduced as compared to LTE for achieving reliability for the uRLLC slice (Bertenyi et al., 2018).

iii. Radio link control protocol (RLC)

This protocol provides error correction, segmentation, and match of transmitted PDU size to radio resources (Bertenyi et al., 2018).

iv. Medium access control (MAC)

MAC provides multiplexing and demultiplexing of radio bearers that are at physical layer. Moreover, it also includes handling the data in priority from radio bearers along with error correction. New functionality is added in NR as compared to LTE MAC protocol. It carries control signaling in a physical layer for the beam-management technology feature (Bertenyi et al., 2018).

b) Control plane

Figure 25 is an illustration of a control-plane protocol stack. Non-Access Stratum (NAS) protocol is used between UE and AMF of 5G core network. This protocol is used for functions related in 5G core such as location updating, authentication, registration, and session management. Radio-related functions, such as control and configuration between UE and NG-RAN, are handled by the Radio Resource Control (RRC) (Bertenyi et al., 2018).

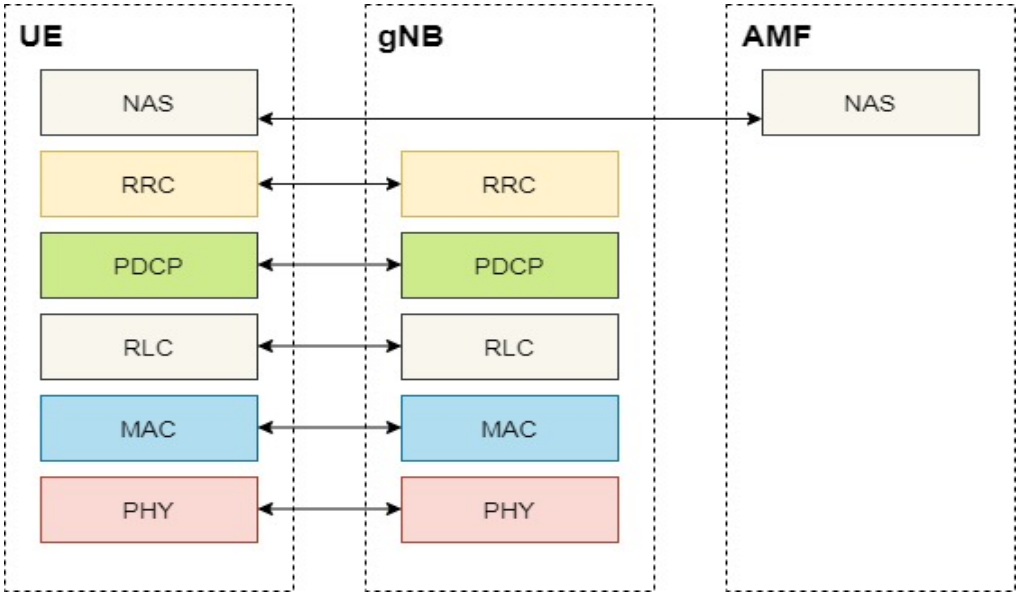

Figure 25. Control plane protocol stack

A 3-state model is introduced in RRC with an additional functionality of INACTIVE state, as compared to LTE RRC. The RRC Inactive, offers a state of battery capacity identical to RRC Idle, but with a UE existing within the NG-RAN; therefore, those transitions to and from RRC Connected are quicker and result in less signaling overhead. The other big changes to LTE RRC are the introduction of an "on-demand" network information mechanism, which allows the UE to request a required precise system information of the NG-RAN wasting radio energy to provide regular periodic system information, as well as the expansion of the measurement monitoring framework to enable beam measurements for a high frequency beam-based deployment (beam forming or Massive-MIMO) (Bertenyi et al., 2018).

c) NG-RAN features

The 5G features are almost overlapping with the 4G LTE Access, with some additional modification in 5G NR. These features and their differences to 4G LTE will be discussed further in this work.

i. Multiple Numerology

Multiple numerology is a subcarrier space that is fixed for LTE, which 15 KHz while in 5G NR there are multiple subcarrier spacings available. In LTE, the duration of a single LTE radio frame is 10ms. The frame is further divided into two slots. One slot of 0.5ms has 6 OFDM symbols in a normal Cyclic Prefix (CP), whereas 7 OFDM symbols used in one slot are an extended Cyclic Prefix. Figure 26 is an illustration of subcarrier spacing (3GPP-TS 38.211, 2019).

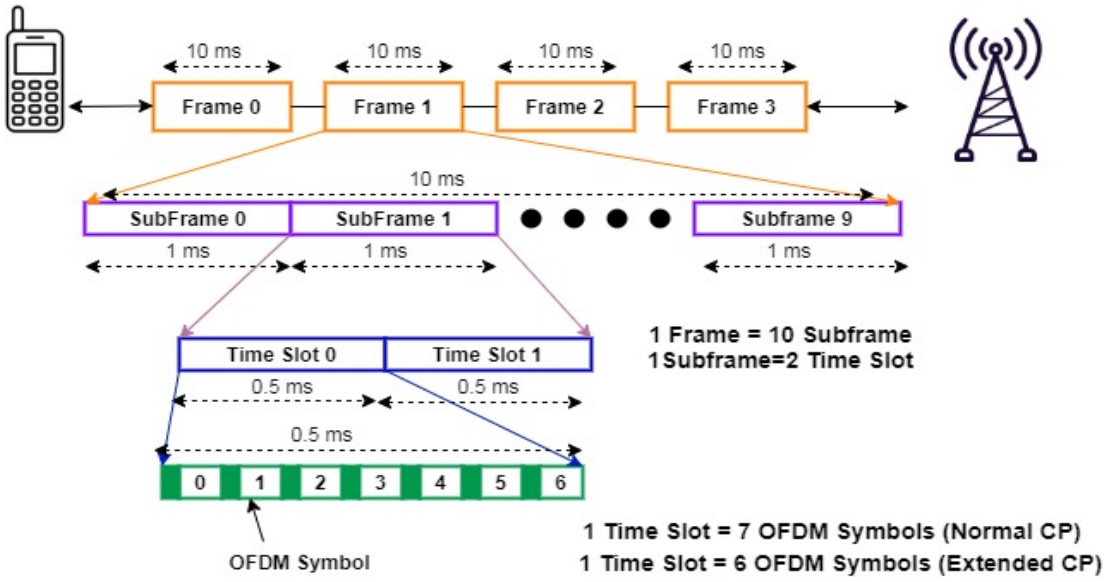

Figure 26 LTE Subcarrier spacing showing Frame and subframe

In 5G NR, subcarrier spacing can be 30 KHz, 60 KHz, 120 KHz and 240 KHz. This becomes notable as the spacing increases the slots per-subframe and while the slot length decreases. The OFDM symbol remains unchanged, that is 6 or 7 per slot, depending on the Cycle Prefix (CP). In LTE, as mentioned earlier, there are two slots per subframe. Figure 27 depicts the carrier spacing for NR (3GPP-TS 38.211, 2019).

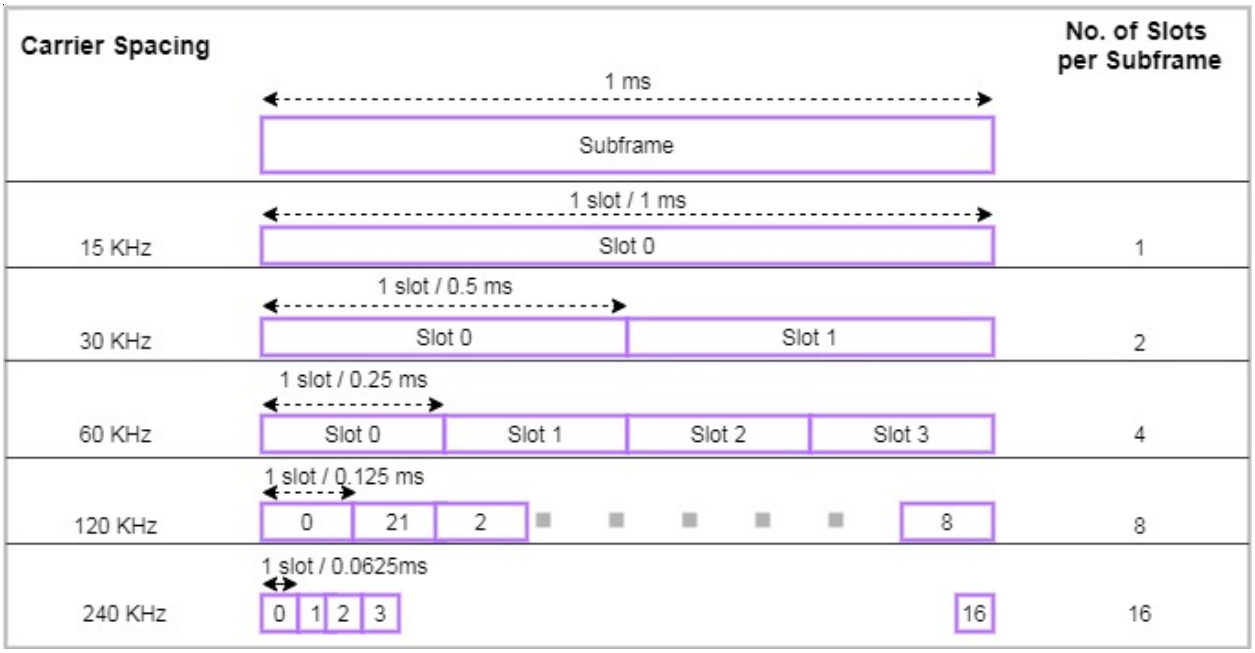

Figure 27. NR carrier spacing

ii. Millimeter-Wave (mmWave)

The Millimeter Wave, also known as high frequency or a very high frequency, has a range from 30GHz to 300GHz. Radio signals are measured by the following wavelength: $\lambda = wave\ speed/frequency, or\ \lambda = v / f$

The wavelength is inversely proportional to frequency, therefore the higher the frequency, the lower the wavelength is, so by converting the 30GHz and 300GHz we get:

30GHz = 9.99308 wavelength = ~ 10 millimeter
300GHz = 0.9993 wavelength = ~ 1 millimeter

According to 3GPPP, frequency range is divided into two categories: FR1 and FR2. FR1 belongs to 6GHz, or "Sub-6" category, while FR2 belongs to 120 GHz sub-carrier spacing. Since wavelength is from 1mm to 10mm, they have high attenuation because of the atmosphere, which can reduce the range and strength of a signal. High objects buildings trees can also block the high frequency waves due to line of sight characteristic of mmWave. To overcome this problem, relaying repeaters should be deployed massively in densely populated areas to use mmWave communication successfully (Harri & Antti, 2009).

iii. Massive MIMO

As discussed in section 1.3.1, MIMO is a wireless technology, which increases the RF radio's data capacity throughput by using multiple transmitting and receiving antennas. According to the 5G goals with massive increase of connected communication devices, it requires a dramatic increase in the throughput of communications as well as reliability. Massive-MIMO is one of the solutions in the use case scenario that will be scalable, easily modifiable and provide reduced communications latency compared to traditional base station antennas.

1.5. Security in Mobile Networks

The number of mobile users will be increased to 13.1 billion by 2023 and onwards with an official launch of the 5G networks (Cisco, 2020). This inclines on the fact that there will be more devices that will share information using the mobile networks, besides the regular mobile users. The mobile operators and telecom industries are becoming the most targeted entities by malicious actors for exploiting the privacy of users and performing espionage on their networks, which can

directly impact the growth of the industry impacted. An example is the attack on Yahoo when more than 1 billion users were impacted due to data breach in a single attack. Another example is an attack on the LinkedIn social networking services, when around 117 million users experienced data breach resulting in private and professional data leakage (McMillan, Knutson, & Seetharaman, 2016). With the requirement for rapid growth of the next generation networks, the network security has gained prominent attention in parallel. Each generation of the mobile networks has diverse security issues, which are explained in the following chapters.

1.5.1. Security Flaws and Protections in 1G, 2G, 3G & 4G

a) First generation Security Flaws

First generation (1G) was based on analog technology also known as Advanced Mobile Phone System (APMS) with different frequencies, using Frequency Division Multiple Access (FDMA) for each call (Liyanage et al., 2018). The mobile network was exposed since there was no built-in authentication to identify the user phone. This led to channel hijacking, cloning and eavesdropping (Liyanage et al., 2018). A cell phone cloning attacker requires hardware and software tools that enables simulating a legitimate network subscriber. At the times of 1G, a radio receiver is used to detect and intercept the user at the mobile tower. The Electronic Serial Number (ESN) and mobile number are obtained from the radio receiver with the help of software, to burn ESN and clone a copy user cell phone by an attacker. This incurs difficulties for the operator to distinguish between actual user phone and cloned phone. The “Oki 900” tool was used by attackers to impersonate and eavesdrop on communication over 1G infrastructure (Liyanage et al., 2018).

Such security threats lead to introduction of the identification mechanism using pin code associated to customer, in order to remove cloned mobile devices in the network (Liyanage et al., 2018). Figure 28 illustrates a cell phone cloning attack.

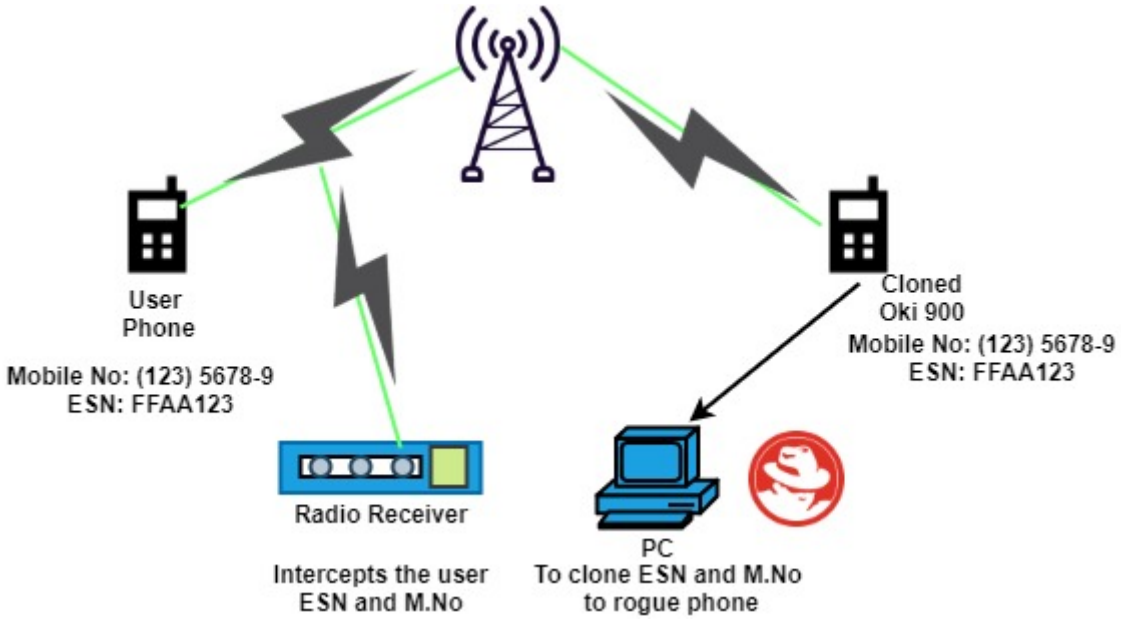

Figure 28. Cell phone cloning attack in 1G network (Liyanage et al., 2018)

b) Second generation security flaws

i. Unidirectional authentication and vulnerability to the man-in-the-middle attack

Authentication was a key focus during the development of the second generation (2G) network standard. The goal was to reduce the mobile cloning attack and channel hijacking. A Subscriber Identity Module (SIM) was introduced as a secure chip inside the phone, using unique identification number. During this phase of protecting the user phone, a new attack has transpired in 2G, masquerading as a carrier base-station with a rogue base-station (Chouchane, Rekhis, & Boudriga, 2009). Rogue base-station also known as “IMSI catcher”, which was used to attack the base station. In other words, to perform Man-in-the-Middle attack (MitM), the two-way authentication was exploited between the mobile user and network by impersonating a trusted Base Transceiver Station (BTS) with rouge BTS (attacker) and allowing the mobile user to communicate over insecure channel with complete visibility of the identity, such as IMSI. Moreover, the untrusted BTS was used to sniff the voice traffic over GSM. In case of GPRS and EDGE, attackers can easily extract the information, including passwords, through user internet traffic. The factor that allows attackers to install rogue BTS is GSM unidirectional authentication. The Mobile Station(MS) with unique IMSI number needs to prove to the Public Land Mobile Network (PLMN) that it belongs to a trusted subscriber but BTS does not require to prove to MS that it belongs to trusted PLMN, which is known by the SIM (Liyanage et al., 2018).

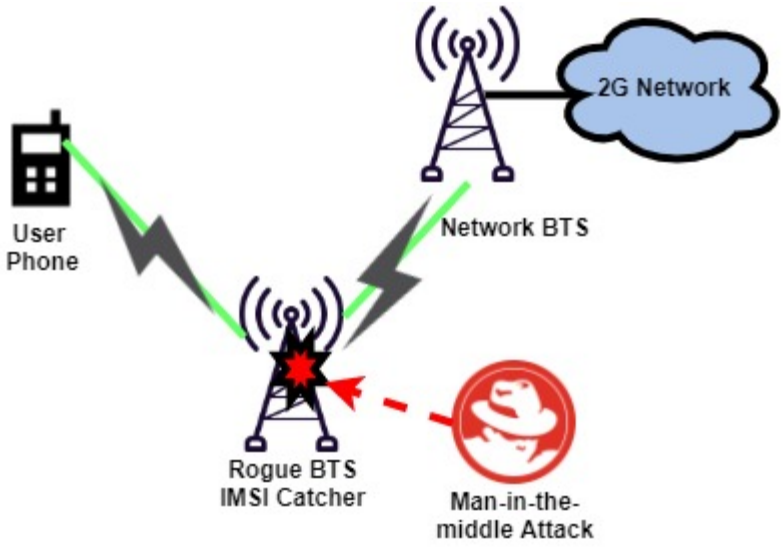

Figure 29. Rogue BTS and IMSIcatcher (Liyanage et al., 2018)

ii. A3/A8 algorithms implementation and drawbacks

The second security threat in GSM is the flaw of implementation of the A3/A8 algorithms. The process of authentication of users in GSM is through challenge-response mechanism. This mechanism involves security components such as random number (RAND), which is 128-bit generated by the Authentication Center (AuC), as shown earlier in Figure 3, a Signed Response (SRES) 32-bit to RAND using authentication algorithm A3 of an individual subscriber authentication key (Ki). The SIM computes the SRES and RAND as an enhanced security, which does not release any information of Ki during this process at the Mobile Station (MS). On the network side (see Figure 5) of GSM system, AuC checks the value of SRES with the value it received from user equipment (UE) and authenticates user in case it is matched. Otherwise, the connection is terminated between the network and UE with a failure message to the subscriber. It is to be noted that the subscriber authentication key (Ki) is not shared over the radio transmission. Furthermore, the A8 algorithm used by SIM, generates a 64-bit ciphering key (Kc) that encrypts data during the transmission between ME and base station (BTS). Thus, A8 algorithm uses both RAND and Ki key, employed in the authentication process, to generate ciphering key Kc within the SIM. This feature leads to resistance on eavesdropping. The A5 algorithms are used to encrypt and decrypt the data by utilizing the ciphering key Kc and are implemented in the hardware part of UE (Cattaneo, De Maio, & Petrillo, 2013).

Figure 30 is an illustration of A3/A8 Algorithm authentication mechanism and connection establishment. The COMP128 is an example of A3 algorithm implementation used for authentication.

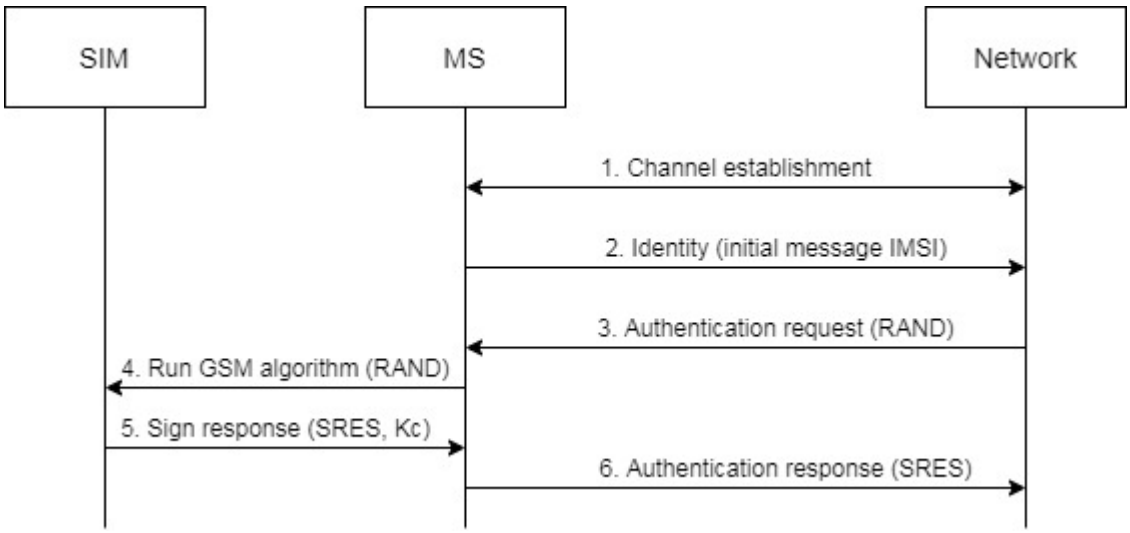

Figure 30. The A3 algorithm process (Liyanage et al., 2018)

Due to the increased processing power of new hardware, the algorithms used for encryption between ME and BTS are vulnerable. Brute force attacks can compromise the A5 algorithm within hours. Although the length of Kc is 64-bit, the last 10-bits are zeros, which makes the key length to 54 bits. Nevertheless, they can still be decrypted with high processing hardware. As a response, GSM has introduced standards for A5 named A5/1, A5/2, etc., which provides an efficient security against brute force attacks, namely sizes of 2^{54} and 2^{16} respectively but it still does not guarantee absolute resistance to cracking. Since the security architecture of GSM is inflexible, it is not easy to increase the key length and thus the same A5 algorithm for encryption is used (Chandra, 2005).

iii. SIM cloning

The third security threat in GSM network is SIM cloning. In case Ki is known, the attacker can easily listen and issue calls, which in return is invoiced to the actual SIM subscriber. Several challenges are sent as an input RAND to the SIM. The A8 algorithm responds then to these challenges as SRES. It is evident that a brute force with an order of 2^{128} can easily provide the information of Ki. In second round of the algorithm, a narrow pipe is isolated into two groups

that consists of 4 bytes. Two bytes are related to Ki and RAND, respectively. This means that attacks with repeated 2R can crack into SIM in 2^{17} RANDS (Singh, Ruhl, & Lindskog, 2013).

Additionally, the "Dejan Kaljevic Sim Scan" is a utility that is used to extract the bytes of Ki with an attempt of 18,000 RAND values. There are two ways for cloning: one is either by physical or the second is over-the-air (OTA). Cloning of SIM is thus a simple process. The attacker must have a card reader and a software tool (e.g. Dejan Kaljevic Sim scan). With weaknesses of the COMP128 algorithm, an attacker can successfully retrieve the Ki and IMSI information by brute force attack as describe previously in the A3/A8 algorithm flaws. Attackers can use Wafer cards (cards that can be programmed with IMSI and Ki information i.e. empty cards) that are easily available in markets and insert the SIM card to extract the IMSI and Ki. These values, once extracted, are then used to create erasable programmable read only memory (EPROM) and Peripheral Interface Controller (PIC) hex files to program the Wafer card. A Personal identification number (PIN) and PIN unlocked key (PUK) numbers are added to access the SIM in the user equipment (Mobile phone). This Wafer card (cloned SIM) is authenticated to GSM network on start of the mobile phone. Consequently, this will allow HLR to generate SRES and K_c using Ki and RAND (from A3 and A8 algorithm), as shown previously in figure 42. Since the length of Kc is 54-bit, a brute force attack will be performed by the attacker by generating RAND (2R, 3R, 4R). The complexity of Kc is 2^{54} by brute force attack but can be considerably reduced with tabulation attack method to obtain Kc within minutes. This similar approach can be used to extract the Ki and with IMSI, a cloned SIM can be reformed to a wafer card (Singh et al., 2013).

In OTA attacks, the SIM cloning is almost like the physical method with an exception that the attacker must eavesdrop over the air interface. In this case, the attacker requires only Ki and IMSI information, apart from Wafer card, as well as EPROM and PIC hex files. The IMSI catcher (Chouchane et al., 2009) is used as a rouge base station between the mobile user and trusted mobile base station (BTS) (see Figure 29). As explained in unidirectional authentication and vulnerability to the man-in-the-middle attack section, this is thereby classified as a man-in-the-middle attack. With the same techniques defined by Dejan Kaljevic 2R, 3R, 4R or 5R, the Ki value can be easily extracted within one hour by just ~18000 RAND responses. Integrated Circuit Card identifier (ICCID) is thus required in OTA, to create EPROM and PIC hex files for cloning as the EPROM contains Elementary Files (EF) of IMSI and ICCID. EF files are being used by

MS and AuC in the Network subsystem. Same ICCID can be retrieved from the mobile applications such as games, anti-virus software which has gained permission to access in the phone. Others applications such as Android protection Antivirus and Webroot have 16 million users ICCID and IMSI information related to the users (Singh et al., 2013). The same can be exploited by attackers to retrieve information necessary for SIM cloning (Singh et al., 2013). Figure 44 illustrates a workflow between rogue BTS and MS with 'N' times of RAND.

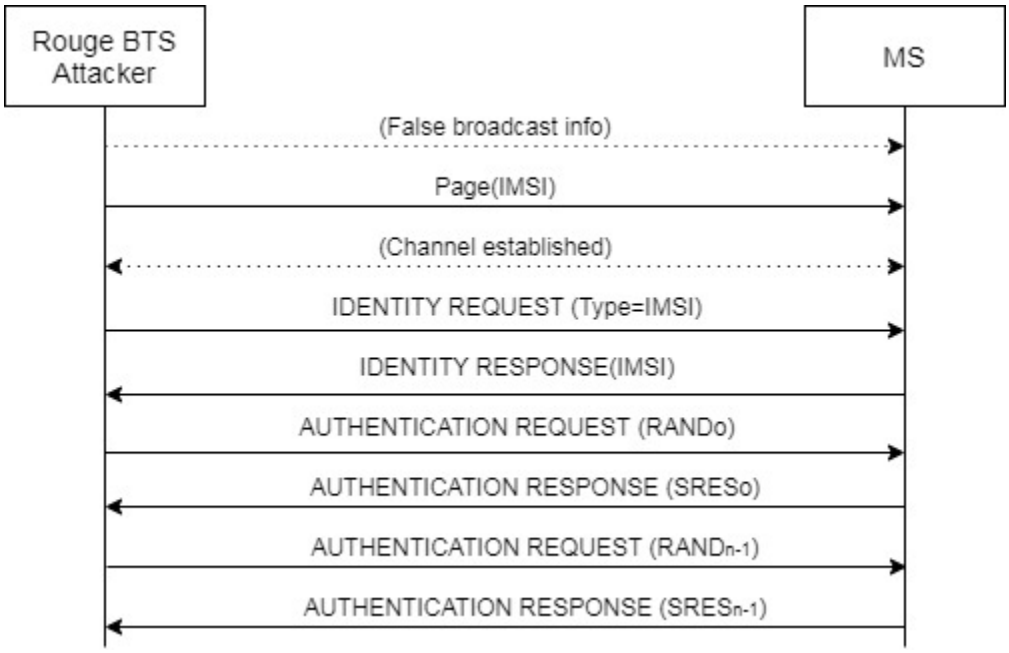

Figure 31. Brute force attack for SIM cloning over OTA (Liyanage et al., 2018)

c) Security in UMTS Environment

The first generation 1G technology, which was based on analog networks, did not have the concept of security features for having a secure connection. GSM networks do have authenticity, confidentiality and anonymity, in which A3, A5 (Eberspacher et al., 2001) and A8 algorithms (Gsma, 2019d) are used respectively. Encrypted data is authenticated by the BTS (Eberspacher et al., 2001) and MS use the encryption by applying the key Kc. A drawback is that the confidentiality exists between BTS and MS (see Figure 32) but not in the entire GSM network. Moreover, A5 is weak encryption algorithm. The TMSI[18] also can be changed by VLR at any time, which may compromise the anonymity as well (Eberspacher et al., 2001).

[18] TMSI Temporary Mobile subscriber Identity allocated to UE once the radio link is established with network. There is no more IMSI exchange over the network after first exchange between UE and network to avoid from IMSI catcher (ETSI-TS 23.236, 2004).

Security has been enhanced in UMTS by addressing the limitations that were known in GSM networks. Below are security features of UMTS:

- **Mutual Authentication:** Itis a procedure in 3G network systems between UE and the network. The procedure includes:

  1) Checking identity and whether MS is acceptable or not,
  2) Providing set of parameters enabling mobile stations to calculate ciphering key,
  3) Calculation of integrity key for enabling the mobile stations,
  4) Permit the MS to authenticate the network (Bais, T. Penzhorn, & Palensky, 2006; Masrom & Ali, 2010).

Figure 32 describes the defined security groups, defined as:

Network access security (I): This feature provides secure access to 3G services and protects from attacks on the radio access link from impersonating fake BTS or node (3GPP-TS 33.102, 2012).

Network domain security (II): These sets of features to exchange signaling data within the enabled nodes MAPsec or IPsec (3GPP-TS 33.102, 2012).

User domain security (III): A Set of features, which secure access to MS (3GPP-TS 33.102, 2012).

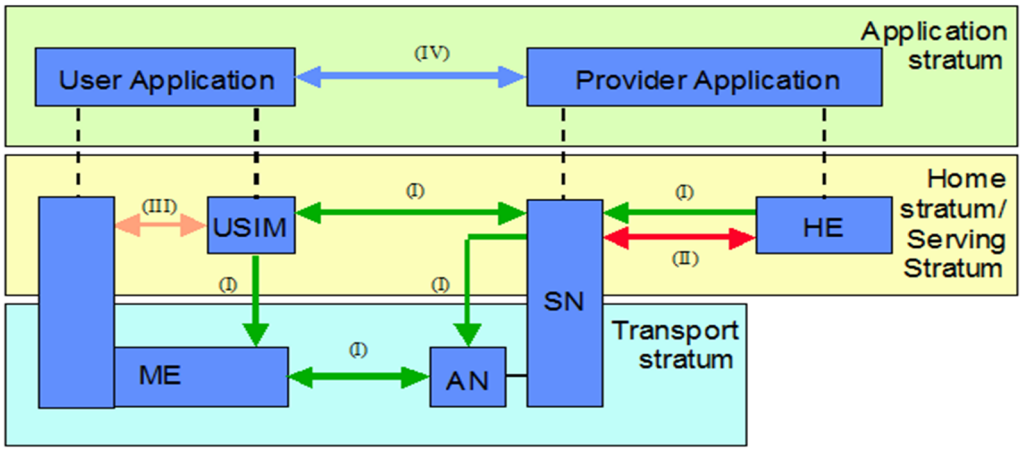

Figure 32. UMTS security architecture (courtesy of 3GPP) (3GPP-TS 33.102, 2012)

Application domain security (IV): Security features that enable applications in the user and provider domains, for secure exchange of messages (3GPP-TS 33.102, 2012).

- Longer key length is used, a 128-bits Authentication and Key Agreement (AKA) as against 64 bits in GSM,
- Security is based within RNC rather on the BS.

While GSM flaws are addressed in UMTS, the security solutions also include the cloning resistance in Universal Subscriber Identity Module (USIM) (Singh et al., 2013). USIM uses MILENAGE and KASUMI algorithms to protect from attackers. As in case of challenge-response communication and generating of Ki and Kc, MILENAGE algorithm creates values during authentication that are responsible in generating the keys. This algorithm consists of values F1, F2, F3, F4 and F5. The input parameters are described as follows:

- **F1** computes authentication token (MAC and MAC-2). These tokens are used to authenticate the network and user equipment with root key K. Input parameters are RAND (128-bits), K (128-bits), a Sequence Number (SQN 48-bit) and AMF value for other specific purpose of implementation (Singh et al., 2013),
- **F2** is a parameter that computes XRES. This is a response similar to SRES (used in GSM) with 128-bits in length (Singh et al., 2013),
- **F3** is a parameter that computes 128-bit ciphering key CK. Inputs values are K and RAND (Singh et al., 2013),
- **F4** provides the integrity key IK used to sign radio control messages. Input values are K and RAND (Singh et al., 2013),
- **F5** is value that computes Anonymity Key (AK). Figure 33 shows the authentication vectors with SQN and RAND parameters as an input for the protection from USIM attacks.

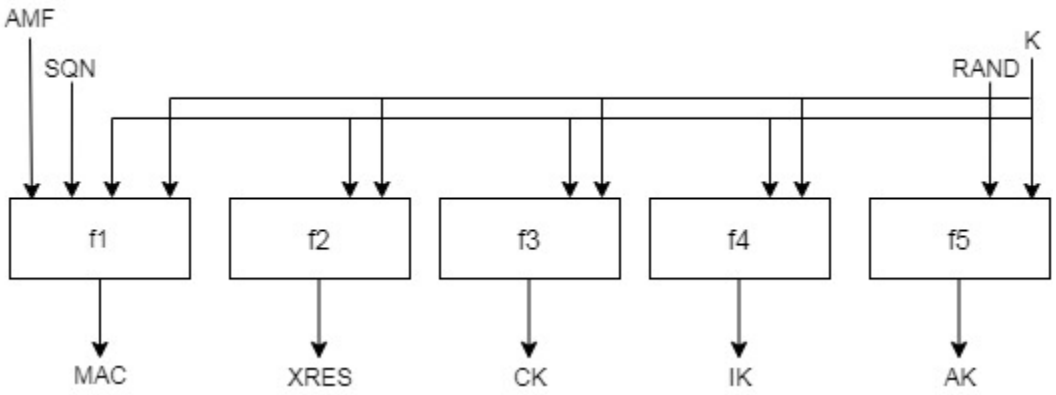

Figure 33. Defending against cloning USIM (courtesy of 3GPP) (3GPP-TS 33.103, 2001)

Authentication token AUTN is generated by the MILENAGE algorithm, which is a replacement of the A3/A8 algorithms to resist attacks in UMTS. Once the authentication is completed, KASUMI (A5/3) is used for confidentiality protection and encryption. This approach of using algorithms provides a strong security as compared to A5/1 and A5/2 used in GSM. With the combination of F values and AUTN, the enhancement has successfully stopped attackers from cloning USIMs OTA to some extent (Singh et al., 2013). Furthermore, as shown in Figure 33, AuC computes AUTN by using F1-5 with SQN numbers, to authenticate the Mobile station. Same process is being executed at the Mobile station with each time AUTN increasing after a successful authentication. The attacker cannot eavesdrop since the SQN number of MS and network correspond to a different SQN, and the authentication will be rejected. Nevertheless, this cannot stop Denial of Service attacks (DoS) in UMTS when the USIM uses the network without mutual authentication e.g. like in the GSM network (Singh et al., 2013).

d) LTE security

LTE security vulnerabilities have imposed many risks to privacy and service availability, and it involves the whole Evolved Packet System (EPS). The classification of LTE security vulnerabilities is shown in Figure 34, which includes both radio access and core networks of an EPS.

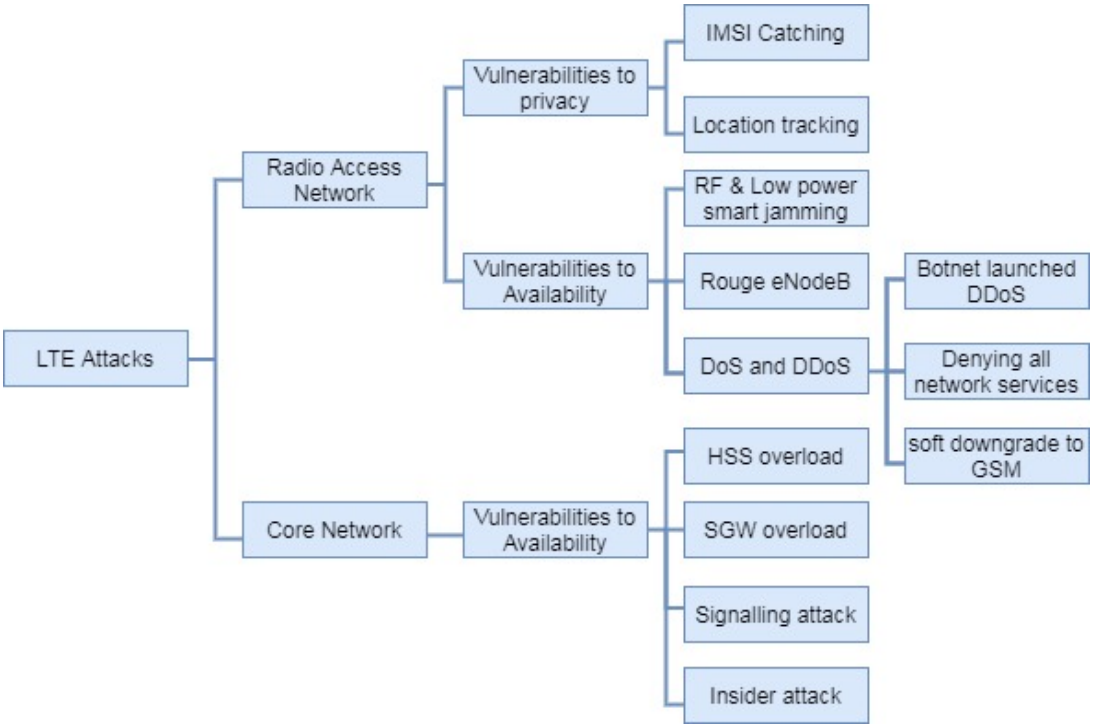

Figure 34. Classification of LTE attacks (Thampi, Madria, Wang, Rawat, & Calero, 2019)

User confidentiality in LTE is vulnerable in various ways. Tracking a user with IMSI ID is leaked through IMSI catcher at radio side of the RAN. Installing rouge eNodeBs, jamming the radio network based on Denial of service (DoS) as well as Distributed Denial of service (DDoS) attacks. DoS and DDoS can incur huge losses for operators, if impacts the core network and can degrade the services availability for verticals and enterprises. Since LTE is IP-based, voice and streaming services are exposed to hackers by installing malware via the mobile data network (Thampi et al., 2019). The attacks in LTE can be classified as: Attacks at the Radio Access Networks and Attacks on the Core Network.

i. Attacks at Radio Access Network

IMSI catching

The IMSI number, a unique identifier of UE, is required for the very first instance when UE is switched on. This process is an attempt to connect with the network. On the other hand, in case there is loss of connection between UE and network, the mobile device must expose the IMSI when the network does not retrieve any TMSI information from user equipment for some interval (Thampi et al., 2019). During this pre-authentication and encryption procedure, the process of

transmitting the IMSI leverage the IMSI catcher to exploit the network (Henrydoss & Boult, 2014).

There can be two viable approaches to achieve IMSI catching attack, either by passive or active methods. In passive approach, the attacker will eavesdrop the traffic over OTA and decrypt all the messages while storing the IMSI information. The other approach is to establish rouge eNodeB, which will ask each device to identify themselves and in return expose the IMSI identity. Moreover, it will give the information including location and conversation messages of users. Femtocells are easily available on the market and are used by attackers for IMSI catching (Thampi et al., 2019).

Location tracking

Location of mobile users is a subject under privacy regulations. The mobile standards have allowed telecommunication industries to extract the physical location of a mobile user. This service can be easily exploited by attackers with passive or active attacks, as described previously. Using trilateration technique which uses distance parameter(Satellite broadcasts signal to get distance) not angle, a mobile user can easily be located within the range of base station (Shaik, Borgaonkar, Asokan, Niemi, & Seifert, 2016).

RF and Low Power smart jamming

As with the case of GSM and UMTS, jamming attacks can occur in LTE as well. This can result in unresponsive network to UE. RF jammers reduce the Signal-to-Noise Ratio (SNR) of the signal causing DoS. In GSM, flood of text messages leads to RF jamming. These text messages share resources with the control signaling channels (see section 1.1.6 for channels description) in GSM and legacy networks. In LTE, the resources are not shared with these signaling channels, and therefore it is not feasible to have text messages flooding in radio access networks (RAN). Smart jamming can be achieved by saturating one or more downlink and uplink control channels, required for the user to access spectrum. Physical control Format Indicator Channel (PCFICH) is more vulnerable to attackers, since it carries the information for decoding the Physical Downlink Control Channel (PDCCH). Traffic and signaling channel information in LTE is known before handshaking that leads to smart jamming. Attacker blocks the downlink channel, which collects the information from PDCCH of UE by using commercial radio jammer targeting frequency band of LTE (Thampi et al., 2019).

Rogue Base station in LTE

In GSM, attacker establishes rogue base station with signal power higher than the trusted base station in range (Vohra, Dubey, & Vachhhani, 2016). In LTE however, the UE will not scan the neighboring eNodeBs because of power efficiency utilization and saving. The "absolute priority based cell reselection" (ETSI-TS 36.304, 2014) can be used to install rogue eNodeB, in which case UE with highest priority frequency will establish connection with the eNodeB (the state of UE in this scenario is IDLE). The attacker will create a connection of UE to the rouge eNodeB, operating with higher frequency cell reselection, despite the UE is located near the trusted eNodeB. The attacker then performs a passive attack to sniff the priority reselection cell list, which is being broadcasted between the real eNodeB and UE to configure rouge eNodeB. With this, it is feasible to achieve Man-In-The-Middle (MiTM), DoS attacks, crashing mobile services and downgrading the LTE network in general.

Botnet for DDoS Attack

Additionally, a Botnet of mobile devices can be used to flood malicious traffic and thus launch a DDoS attack. As a result, this will obstruct the links to downgrade the network, while denying services to mobile users. This malicious traffic impacts both RAN and the core of the network (Thampi et al., 2019).

Another example of DoS attack technique is plain reject messages (Shaik et al., 2016) that will allow a rouge eNodeB to reject UE from accessing LTE services. This reject message does not require mutual authentication as per LTE specification (Thampi et al., 2019). When UE is connected to rouge eNodeB, it sends “Tracking Area Update (TAU) Request”, which is integrity-protected but not encrypted. The attacker decodes the TAU message and replies with “TAU Reject” message, without integrity protection. Once the "TAU Reject" is accepted by UE, it will remove LTE services that are part of real network. To activate the LTE services again, USIM needs to be reinserted or the UE rebooted. However, in meantime, the UE will search services from other legacy networks such as GSM or UMTS. Due to downgrade of services to either 2G or 3G, attackers will have more potential to exploit the UE in the form of MitM attack, eavesdropping to calls and text messages, etc. Figure 35 illustrates the Downgrade to non-LTE services.

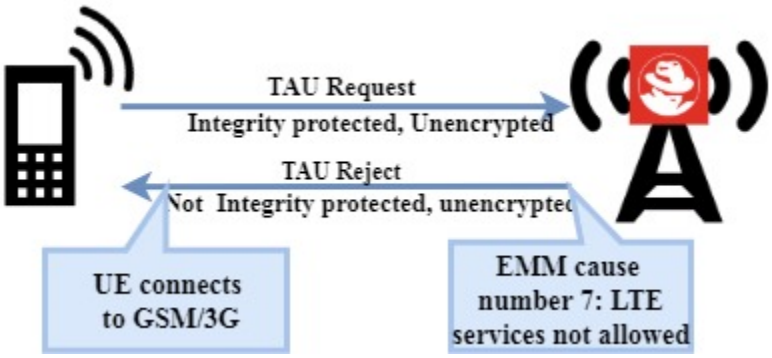

Figure 35. Downgrade to non-LTE services (Thampi et al., 2019)

ii. LTE core attacks

The previous examples conclude that DoS/DDoS are serious threats to LTE network elements, which can affect Radio Access Network (RAN) or the Core Network (CN). Mobile botnets can utilize maximum resources of the network, resulting in unavailability of services and network congestion. These flood of messages can have impact on core network during attach/detach procedure for Serving gateway S-GW, PDN-GW and MME (Thampi et al., 2019). Apart from this, it can also affect Home Subscriber Server (HSS) where the user information resides.

Home subscriber server Overload

The HSS stores information such as IMSI, phone numbers, authentication keys and last known location of UE (see section 1.3.2 for more detail). An accomplished attempt of DoS/DDoS overloads the HSS by continuously sending false IMSIs. In response authentication, vectors are being generated to establish the connection. The repetitive response of authentication vectors, generated by requests, utilizes excessive resources and thus resulting in HSS overload (Thampi et al., 2019).

S-GW saturation

Serving Gateway (S-GW) overload can also prompt unavailability of LTE network services. Bearer and triggering Tracking Area Update (TAU) flood messages can saturate the S-GW. As shown in Figure 35, during the signaling procedure, MME implements security parameters for TAU. These parameters include integrity check and REQUEST message. Based on these parameters, guarantees real network users. The problem with this implementation by MME, is that it takes place before user authentication in LTE network occurs and MME assumes that the user is trusted and will not further attempt an authentication process until the integrity is broken.

Due to this fact, DoS/DDoS is possible to be performed on MME by attackers. Whenever the UE moves to new location tracking area, MME(compromised by attacker) can send floods of bearer messages resulting in S-GW overload (He, Yan, & Atiquzzaman, 2018; Thampi et al., 2019).

Insider Attack

These attacks are not prominent and are less likely to occur. An inside attacker has a privilege access to core elements of LTE network and can shut down a eNodeB, HSS, MME or S-GWs, or even launch botnet of mobile devices. The outcome of this attack can globally effect the services If launched on scale (He et al., 2018).

e) 5G security challenges and solutions

Diversity of 5G services and applications, such as eMBB, URLLC and mMTC, are prone to threats and contain multitudes of purposes to attack 5G services, compared with previous network generations. There are numerous possible vectors in 5G that varies from user equipment such as mobile phones, Robots, Internet of Things (IoTs), automated industrial equipment, autonomous cars, etc. Figure 36 illustrates a 5G security threat landscape, which includes Radio Access Network (RAN), core and the Internet. Different threat types and possibilities of area of attack are shown, which are Man-in-the-middle (MitM) attacks at Cloud RAN (C-RAN) domain, IP core networks being prone to DDoS and user equipment being exploited by ransomware, malware and Bots (Liyanage et al., 2018).

The transport network slice segment resides in the Cloud RAN and the vulnerabilities of the SDN controller directly translate to the transport network slice. It is necessary for mobile operators to consider the security challenges related to them since the whole infrastructure of 5G mobile networks depends on them including network slice. If anyone the technologies are vulnerable to attackers, the whole infrastructure can collapse. 4G and legacy mobile networks were not designed to cope the security threats related to NFV and SDN. The need of security is required to both signaling and data traffic at multiple points of attacks (as shown in Figure 36) from user equipment's (UE's) to RAN and core network elements(Liyanage et al., 2018).

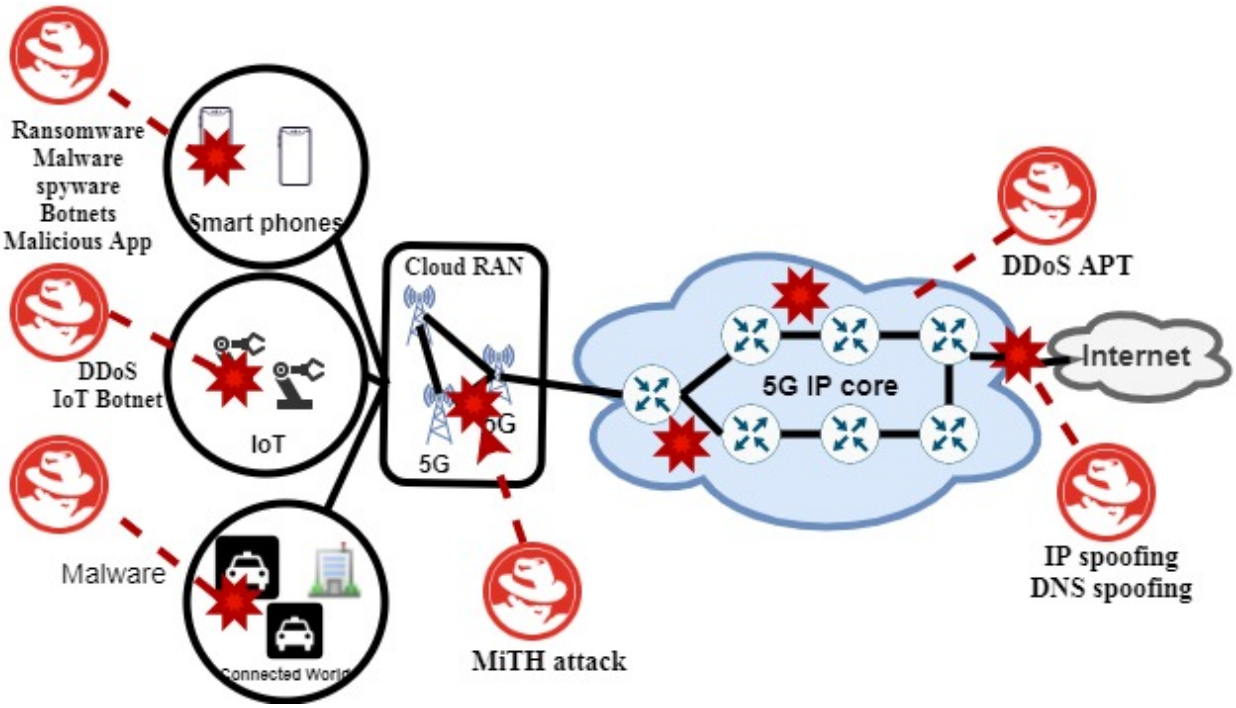

Figure 36. 5G threat landscape (Thampi et al., 2019)

i. SDN security challenges and solutions

SDN separates the data-plane and user-plane, while providing programmability and a logically centralized control-plane. This is a prime feature in realization of network slicing and 5G mobile networks in general, but the very same can be vulnerable to security threats. The SDN controller updates the flow rules to the forwarding plane nodes, due to which this update information is identifiable in a network and can lead to DoS attacks. Moreover, in some instances centralized controller is required that can result in creating a bottleneck for the operator network in saturation attack scenario. In the case of a transport segment of the slice, the Traffic Provisioning Manager (TPM) and Customer Network Controller (in ACTN framework) grant an access to external applications to create the enhanced VPN tunnel, but this access can be exposed to malicious application through insider attacker in the customer's organization. Scenario like this can be very destructive inside the operator's network (Ijaz et al., 2018; Y. Lee & Kaippallimalil, 2019). If the SDN control-plane is not resilient to security attacks, it can affect the data-plane forwarding elements that are subjected to saturation attack vectors. This is because the OpenFlow switches have a limited unsolicited flows of TCP/UDP packets (Ijaz et al., 2018). Third possible attack related to SDN is payload-based attack. In a presumed scenario, a stateful firewall, Intrusion Detection System (IDS) and Intrusion Prevention System (IPS) are deployed at application layer

as a service. These applications scan the packets individually and the controller receives instructions to follow the rule for in-depth packet inspection by IDS. For instance, the first packet's header does not provide sufficient data if it is suspicious, later allowing malicious traffic by being considered as a trusted packet. In this way, the controller will not be able to detect the payload of packets that cause a threat to the network (Khettab, Bagaa, Cadette Dutra, Taleb, & Toumi, 2018).

(Khettab et al., 2018) has used the SECaaS framework to detect the malicious flow of traffic. They have proposed that once the malicious traffic is detected, it alerts the security orchestrator to command the controller to stop the traffic and reduce the bandwidth for malicious traffic flow, in order to avoid any degradation of services. SECaaS framework consists of deep packet inspection DPI engine to prevent payload-based attacks.

ii. NFV security challenges and solutions

Network Function Virtualization in 5G will provide means to install network functions as required, without installing a new hardware equipment. It is a way towards agile networking and certainly a cost-efficient solution for telecommunications. Having a positive approach in deploying these functions, there are important security concerns in the underlying infrastructure of networks as these can impact system resiliency and thus can effect quality of service QoS (Taleb, Ksentini, & Sericola, 2016). It is evident that the hypervisors were vulnerable previously to security attacks, for example VM/guest OS manipulation and data exfiltration/destruction. Due to these, other components are vulnerable to attacks in parallel. Moreover, NFV enables an automation of provisioning network functions through orchestration, which can lead thus to orchestration exploits, malicious misconfigurations and SDN controller exploits. Interoperability is yet another issue that can cause a security loophole in infrastructure, since vendors have different VNFs solutions (Yang & Fung, 2016). Following are some security risks related to NFV.

Isolation Failure

This attack is deliberated when an attacker manages to compromise the hypervisor on which VNFs are running. It is also called VM escape attack. Figure 49 shows an illustration of VM escape attack in steps.

Step 1:VNF is compromised and gaining access to operating system becomes possible. This is a result to improper isolation of hypervisor and VNFs.

Step2: Attacker gains access to VNF connectivity using tools with management APIs. (Launches application that can send packets and result in overflow).

Step3: Attacker accesses hypervisor to cause an impact. (Code to gain access to the host through hypervisor) (Lal, Taleb, & Dutta, 2017).

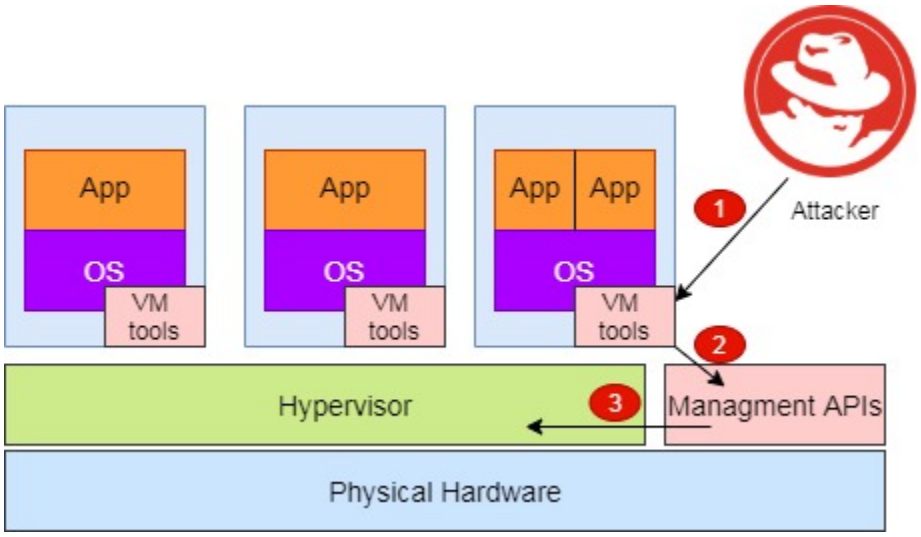

Figure 37. VM escape attack (Lal et al., 2017)

DoS amplification attacks

Service availability can be impacted if VNFs are exposed to DoS attacks. Huge burst of traffic can be generated from compromised VNFs that is being run on the same hypervisor, which will result in affecting other VNFs. Another possibility is Domain Name Server (DNS) amplification attack in which attacker performs two malicious activities. Initially attacker will spoof IP address of DNS resolver and replacing with attackers IP address Replacing of DNS resolver IP address with attacker will send replies to attacker's server. Secondly the attacker gets the Internet domain which is registered with other DNS records. The attacker will send DNS query which in return request all the DNS records for specific Internet domain. These queries in result perform large DNS responses from DNS servers thus resulting amplification effect (Lal et al., 2017).

Malicious Insider attacks

These attacks are carried out by an administrator who has a privilege access to network infrastructure (or someone gaining a privileged access to the same). The administrator goes through a process of memory dump of user VMs and can extract information, such as passwords, usernames and SSH keys. This violates privacy and data confidentiality. Additionally, administrators of 5G cloud infrastructure can be threat in the form of internal attackers by extracting the information from hard drives used by the cloud infrastructure. A backup copy is thus created and tools such as kpartx and vgscan are used to extract information (Rocha & Correia, 2011).

NFV threat solution and summary

System components (OS kernel, BIOS) can be secured by using trusted platform module (TPM) as a hardware root. This provides isolated storage and cryptographic identity and is also known as hardware-assisted isolation. Moreover, the hypervisor, as previously shown in Figure 54, provides virtualization between underlying networks and VMs that are communicating with other networks using SDN. The hypervisor patches should be updated regularly and disable unnecessary services that are not in use, such as SSH and remote access. The concept of network slicing can prevent VMs from impacting other VMs by separating traffic, respectively. A demilitarized zone (DMZ) can be utilized to separate traffic into zones. Another approach to secure the VMs is hypervisor introspection, used to find abnormalities in software that runs in VMs. These tools provide deep inspection of VMs, which in result provides powerful method to secure the infrastructure. Table 5 is a summary for NFVI security risks and best approach to secure them (Lal et al., 2017).

Table 5. NFVI security threats and solutions summary (Lal et al., 2017)

	Security risk	**Target**	**Best approach**
1	Compromised hypervisor	Platform	Separation of management traffic, hypervisor patching
2	Isolation failure	Platform/VNFs	hypervisor introspection, security zoning
3	DDoS attack	VNFs	Flexible VNF strategic deployment
4	Malicious insider	VNFs	VNF image signing, operation guidelines

1.5.2. Virtual Private Networks (VPNs) and their role in mobile networking

This section focuses on Virtual private networks (VPNs) and how they are implemented using tunneling technologies in a mobile network. Mobile infrastructures carrying user traffic are exposed to a serious problem in terms of security and privacy protection (Snader, 2005). Private networks are created to address this problem within the infrastructure of mobile and enterprise networks. These private networks (VPNs) utilize tunneling technology that ensures authenticity and privacy protection within a network flow. The data in a tunnel can be encrypted, thus should be exposed to other vertical industry services (Snader, 2005). In other words, private networks provide traffic isolation of users that is a logical network implemented over the shared physical infrastructure of the service provider. (J. Dong, Li, Miyasaka, Bryant, & Young, 2020) define VPN definition as *"A network formed by applying virtualization on public physical network infrastructure, in such a way that the users are able to use it as a private or user-owned network."*

VPN is not a new concept as in the past the ATM/Frame relay technologies were providing VPN services. Nevertheless, later the mobile operators started to deploy IP-based infrastructure, thus opening the possibility for also implementing IP-based VPNs.

The 4G LTE architecture is based on flat all-IP backhaul network, which transports traffic of different types such as the ones from Evolved-NodeBs (eNBs), Service Gateways (SGWs), Mobility Management Entities (MME) and cross-handover traffic on the X2-U and X2-C interfaces between the eNB base stations. Furthermore, in LTE the flat IP architecture distributes Radio Network Controller (RNC) functions with eNBs, MME, S-GW and are directly connected to the core network (Juniper Networks, 2015). This leads to challenges to provide a secure traffic in backhaul networks in LTE (Liyanage & Gurtov, 2012). Other challenges may include routing traffic to proper destinations. Moreover, it should be noted that these may effect quality of service (QoS) and cause various discrepancies in the traffic flow to the end users (Liyanage & Gurtov, 2012). One approach to remedy these obstacles is IPsec VPN architecture, which can solve issues in LTE network (Liyanage & Gurtov, 2012). The three prominent protocols for IPsec are as follows (Snader, 2005):

- Authentication header (AH): protocol that ensures data origin authentication and integrity,
- Encapsulating Security Payload (ESP): protocol that provides data privacy including AH protocol functions,
- Internet Key exchange (IKE): protocol that provide key management functions.

(Snader, 2005) has used IKE protocol for implementation of IPsec that negotiates a key management between the two endpoints to protect a VPN tunnel. IKE protocol is called hybrid protocol as it is the combination of three key-exchange protocols named as Internet Security Association and Key management protocol, Oakley key determination protocol (OAKLEY) and SKEME protocol (Snader, 2005). Authorization server (AS) is used to authenticate the new user accessing the VPN, based on IPsec tunnel and logical interface are separated with unique IPs in order to distinguish the mobile traffic at gateways (end points). This solution proposed is based on Layer-3 IPsec tunnel (Liyanage & Gurtov, 2012). Moreover, IPsec solution was implemented in tunnel mode in this case. IPsec has two types of modes: transport and tunnel mode, respectively. In transport mode, two fixed hosts are connected through VPN and this does not provide connectivity between networks or between the host and the network. The end-to-end security is guaranteed in transport mode, but it has a constraint of network to network connectivity. In tunnel mode, a connectivity is possible between networks and host/network with a requirement of large bandwidth, which is an expensive solution. Gateways are involved to route the traffic along with performing encryption, decryption, and authentication of traffic between two endpoints. While implementing VPN tunnels using IPsec has advantages, on the other side there are disadvantages as well. (Jyothi & Reddy, 2018) describes Quality of Service (QoS) in videoconference and how it is affected by packet loss using IPsec tunnels due to traffic load between the two networks (host to network/network to network). Traffic load is increased because the IPsec used AH and ESP protocols to protect the data which overloads the header of packets. Due to overhead of packet, this increases traffic load and in result can cause traffic congestion in the tunnel and prompt to packet loss (Jyothi & Reddy, 2018). At this point, if the same IPsec tunnel is used as a VPN service in 5G network, it is not feasible solution since packet loss cannot be afforded in case of URLLC or mMTC where we have millions of sensor

communicating and complex machine to machine communication is occurring (Jyothi & Reddy, 2018).

The second solution to distinguish the traffic between the end points in a VPN tunnel is Host Identity Protocol (HIP). HIP works in between Transport and IP layer known as HIP layer and is identified with Host identity Tag (HIT). The real IP address will only be used at the end points in order to reach its proper destination (Liyanage & Gurtov, 2012). HIP implementation is performed based on Internet Protocol version 4 (IPv4), ignoring internet protocol V6 (IPv6), indicates that VPN carries encrypted traffic of IPv4 not IPv6. Ignoring IPv6 traffic in a tunnel can cause VPN traffic leakages (Gont, 2014). Furthermore, sending IPv6 traffic in clear text exposes a vulnerability to inside attackers in breaching the main security protection elements, that are confidentiality and service privacy (Gont, 2014).

Third common approach of backhauling traffic in mobile and enterprise networks is multiprotocol label-switching MPLS-VPN model (Jyothi & Reddy, 2018). MPLS-VPN separates the user traffic with MPLS tag that is based on Layer-3, but it does not provide authentication of end nodes, data encryption or privacy protection. MPLS architecture uses Border Gateway s (BGP) that is used to establish tunnel and employs TCP Layer-4 sessions, which are vulnerable to TCP/IP based attacks(such as TCP DoS, reset attacks) (Liyanage, Kumar, Ylianttila, & Gurtov, 2016). Moreover MPLS an expensive approach which involves many steps that includes Customer Edge (CE) routers, Provider Edge (PE) routers, Provider (P) routers to route MPLS traffic between two VPN end points (Jyothi & Reddy, 2018).

Generic Routing encapsulation (GRE) is a VPN tunneling protocol developed by Cisco systems (Idrissi, Elkamoun, & Hilal, 2019). A GRE tunnel establishes connection between two networks that appear as a homogenous network from the inside. Packets inside the GRE tunnel are not parsed through nodes while being sent to another network. Original data is encapsulated with outer GRE header and routed to destination network node. The original packet is then removed with an outer packet (header information) at the target node in the network. GRE is an efficient VPN solution in terms of QoS but has drawback in implementing since it does not provide encryption, which can add to a security vulnerability. IPsec can be used for authentication and encryption over GRE for data protection, but the performance of data

transmission may be affected as compared to the generic GRE tunneling (Cisco, 2006). All tunneling technologies discussed are based on soft isolation of services (J. Dong et al., 2020).

1.6. OpenAirInterface5G Open-Source Alliance for democratizing 5G implementation

The OpenAirInterface5G (OAI) project by EURECOM is an experimental open source software proto-type environment to provide innovation in mobile communications (EURECOM, 2020). OAI includes functionality of (base stations (eNBs), user equipment (UE) terminal, and core network), OAI-RAN radio access network model and a 5G-Core / Next-Generation NodeB (gNB) functionalities. Furthermore, OAI uses general purpose processor (x86, ARM) for the realization of cellular network functionalities (eNBs, UE, core, Baseband Units, Remote Radio Heads etc.) under the compliance of OAI software Alliance license model. The latest standard Release 10 LTE for eNB, Core is implemented on Linux based computing (Intel x86, ARM) following 3GPP standardization process from release 13 and path towards 5G.

OAI platform has two primary features as follows (Nikaein et al., 2014):

- OAI is an open source implementation of 4G/5G mobile cellular system can be used for real-time indoor/outdoor experiments demonstration,
- Built-in emulation capability that resembles with the real execution of environment.

OAI network installation is divided into three nodes as follows (Dzogovic, Do, Feng, & Do, 2018)

- Commercial enabled mobile equipment (COTS LTE UE),
- LTE core network (OAI soft EPC realization) running on Linus commodity computer,
- OAI soft EPC and eNB replaced with dedicated hardware elements such as MME, HSS, S-GW and baseband units (BBUs).

Figure 38 shows a standard OAI architecture.

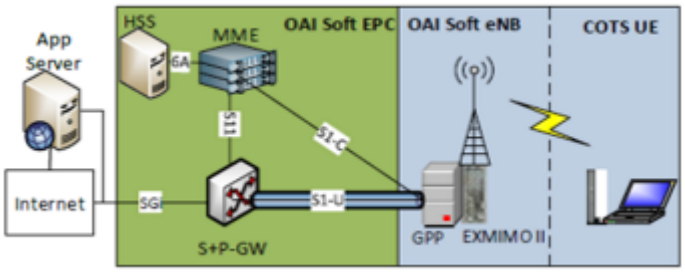

Figure 38. OAI architecture (courtesy B. Dzogovic) (Dzogovic et al., 2018)

(Dzogovic et al., 2018) employed containerization for the deployment of OpenAirinterface by running Evolved Packet Core (EPC), Home Subscriber Server (HSS) and eNB in a single container, building a virtualized 5G network. A container as a virtualized environment, can pack applications that can be deployed instantaneously on single or multiple physical or virtual machines (VM). Each container has its own set of libraries to run the software. Since they run on same host's operating system, they are more efficient that virtual machines in terms of memory, CPU utilization and storage (Docker Inc., 2020a, 2020b). Furthermore, (Dzogovic et al., 2018) described scalability as EPC constituents can be deployed in separate containers. This same applies for the eNB function, which can be split into Baseband Units and Remote Radio Heads, instantiated in distinct containers. To achieve this, Docker is used for implementation of the containerized OpenAirinterface5G, sharing same OS kernel and starting instantly with less CPU and memory utilization. Communication between hosts with EPC interacts with eNB through Layer-3 networking and GRE tunneling, which leads to performance issues such as high latency, but this scenario (L2TP tunneling soft isolation) fits for communication that involves Internet of Things (IoT) for home devices or personal user purposes. In addition to that, (Dzogovic et al., 2018) did not implement IPsec, which means the traffic is unencrypted in the tunnel.

Complementary to the aforementioned, (Feng et al., 2019) has provided an alternative way of implementation of the network slicing concept in which users accessing services are isolated from other slices. The experiment was performed in the Secure 5GIoT Lab at the Oslo Metropolitan University with initial establishment of 5G network, comprised of Cloud Radio Access Network (C-RAN) communicating with a core instantiated in an OpenStack cloud. In this case, GTP-U (GPRS Tunneling protocol User data tunneling) is used in the mobile network for establishing communication to IoT devices and User Equipment (UE).

(Bruno Dzogovic et al., 2019) performed another area of research that can enable a way to provision of high-bandwidth network slices with Thunderbolt-3 networking using concept of virtualized SDN cloud radio access network (C-RANs). Two techniques are used such as the Intel Thunderbolt technology and SR-IOV (Single Root - Input/output virtualization) for functional split of the base station functionality to achieve higher bandwidth between Base Band unit (BBU) and Remote Radio Heads (RRHs). SR-IOV is a way to define the process of virtualizing PCIe (Peripheral Component Interconnect Express) devices. PCIe is a high speed interface standard connecting high speed components (Dhawan, 2005). SR-IOV enables virtualization within PCI which further allows virtual machines (VMs) to share PCI resources means user will achieve latency by bypassing the hypervisor and virtual switch layer (Bruno Dzogovic et al., 2019). Two functions Physical functions (PFs) and virtual functions(VFs) are implemented in SV-IOV (Intel, 2011).

Nevertheless, the experiments do not cover the transport network (TN) between the Network Core and the Baseband Unit, which can be a subject to multitenancy in a shared cloud. Such an environment can prove sensitive for storing vast amounts of personal user information and credentials, especially when the transmission of those is achieved via share virtualized network resources.

2. Network Slicing

5G networks are expected to provide three main services: Massive Machine Type Communication (mMTC), Ultra-Reliable Low Latency Communication (uRLLC) and enhanced mobile broadband (eMBB). Each of these three services are characterize by some requirements. In case of mMTC, reliability is mandatory, whereas low latency and ultra-reliability are required in uRLLC. For eMBB, higher data rates are the goal of 5G, to enable end users better service and user experience, as well as cope with the content demand and higher data flows. For this reason, it has become necessary to redesign the network architecture to allow users and vertical industries to use these services. Network slicing is a key feature to enable this variety of services and applications in 5G networks. This includes E-Health, smart transportation, augmented reality, smart farming and many more. The idea of network slicing is to exploit the network infrastructure to build numerous sub-networks that contain all the necessary services and applications. Furthermore, these sub-networks can exploit the physical shared resources to

provide the services as an isolated network without any caveat during peak hours and may diminish during idle periods to reduce the costs, which indeed is scalable and flexible. This isolation may additionally guarantee the reliability and security of data between the vertical industry applications and allow specific user to access the services securely. It is possible to create a network to enable 5G services that can have fixed physical resources with very high bandwidth to avoid congestion and delay in services and applications, but it is not practical to implement due to both capital expenditure (CAPEX) and operation expenditure (OPEX). It is thereby feasible to allow the sharing of physical resources with an overall end-to-end isolation enabled among multiple slices, utilizing the capabilities of Software-Defined Networking (SDN) to build virtual Network Functions (vNFs) for that purpose (Kazmi et al., 2019; Zhang, 2018).

2.1. Concept and Principles of Network Slicing

The notion of network slicing represents a logical network or networks, situated on a common physical infrastructure also known as “virtual network”. The network softwareization (Afolabi, Taleb, Samdanis, Ksentini, & Flinck, 2018) concept will enable network slicing through software-based solutions. For that purpose, network slicing in 5G will utilize SDN, NFV, cloud computing and edge computing for the realization of network slices over the same physical infrastructure. Each slice will have the independent control and will have ability to scale according to requirements. Figure 39 is an illustration of logical networks of different service types and how network slicing helps enable the same (Ordonez-Lucena et al., 2017).

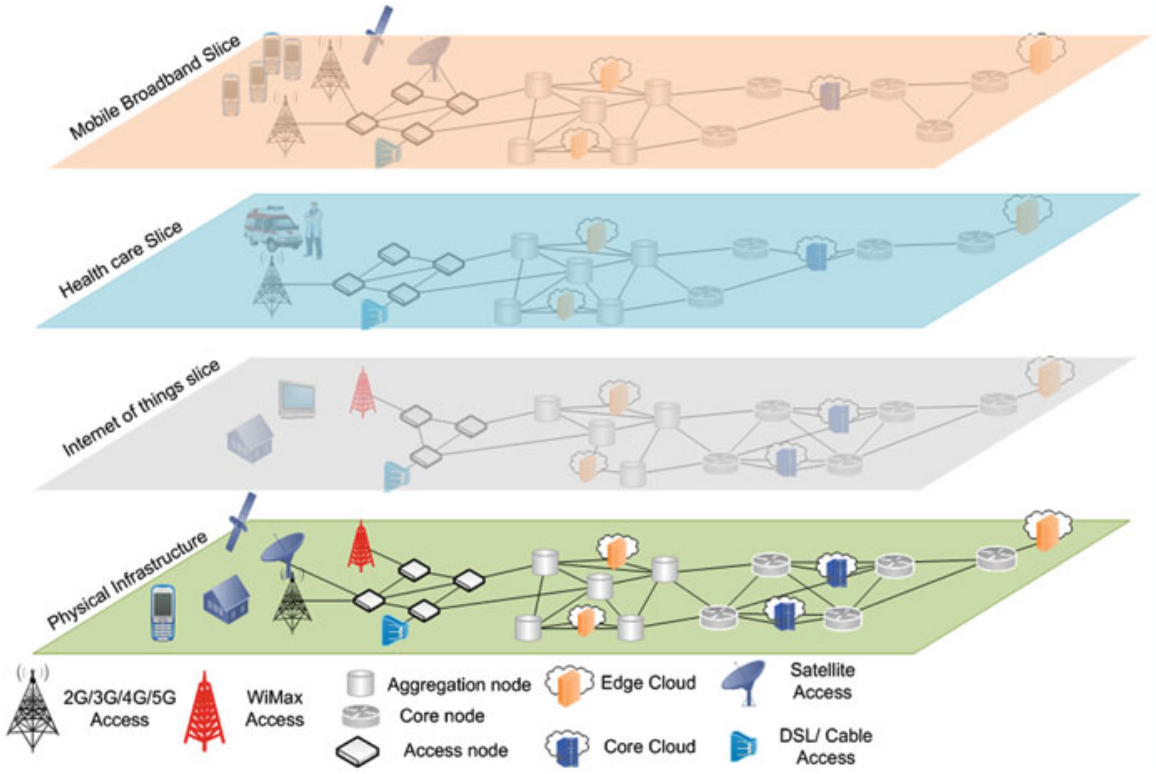

Figure 39. 5G network slicing (courtesy 5GPP) (5GPPP, 2019)

The 5G Infrastructure by Public Private Partnership Project (5GPPP) has proposed network slicing architecture consisting of five layers, such as infrastructure layer, orchestration layer, business function layer and network function layer (5GPPP, 2019). On the other hand, NGMN alliance proposes network slice division in three layers: business application, business enablement and infrastructure resource (NGMN, 2015). According to (Foukas et al., 2017), a three-layer framework for network slice is presented and includes: infrastructure layer, service layer and network function layer (see Figure 40). Management and Orchestration (MANO) perform the operation of creating network slicing as per use case requirements.

The infrastructure layer deals with physical resources that cover the transport segment, core network and radio access network (RAN). Moreover, this layer controls the resources and allocation to slices. In case of emergency use cases, this layer deals with the scalability of slice and isolates the traffic, accordingly. A software implementation of network function is performed in a network function layer, which is used for concatenation or service chaining of network slices.

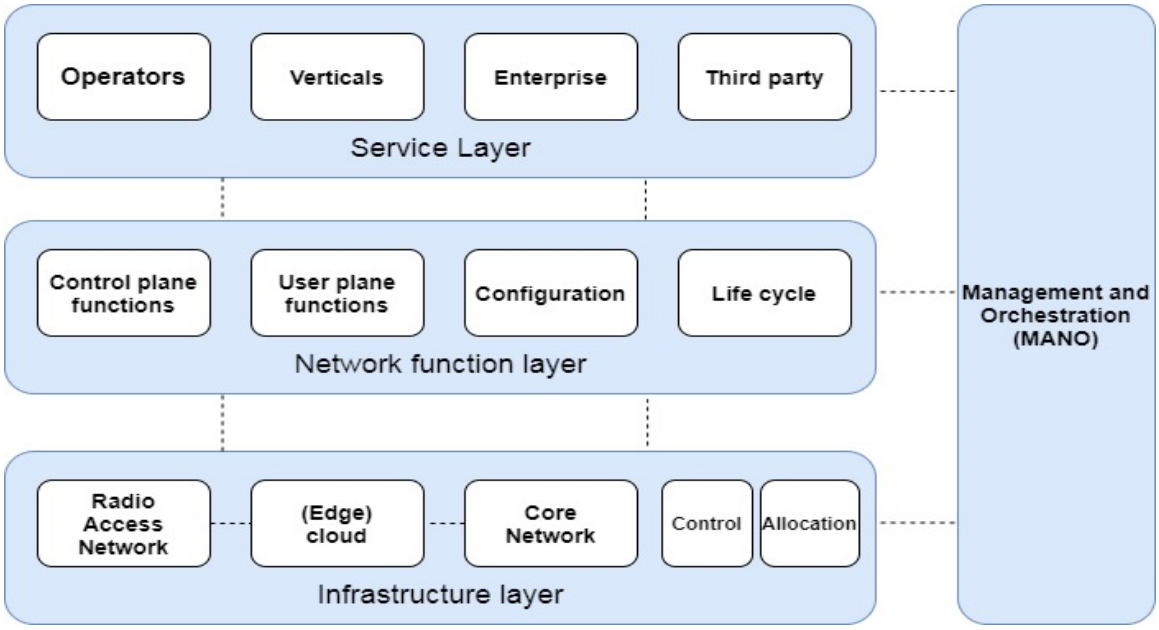

Figure 40. 5G network slice layered framework

a) Software-Defined Networking (SDN)

SDN provides separation between control-plane and user-plane functionalities. The network is divided into two planes: a data transmitted through the data-plane as well as a control-plane to provide centralized control layer functionalities. Therefore, it becomes easier to manage the network due to the centralized nature in SDN (Kazmi et al., 2019).

b) Network Function Virtualization (NFV)

NFV provides the infrastructure, which is comprised of functionalities of the network nodes that are virtualized, and further concatenated to provide different network services. This virtualized network may have virtual machines, virtual servers, virtual firewalls, edge computer nodes and cloud computing infrastructure. Consequently, these may provide dedicated or dynamic resource allocation based on their configuration.

c) Tenant

The tenant is an individual user or vertical industry entity that can access shared resources and has been granted privilege rights to access the same.

To achieve a secure and scalable network slice, particularly for the use case in an emergency services scenario where a transport segment slice has to be agile with performance guarantee,

there should be a specific framework. Since the goal is to achieve complete isolation for secure communication, integrity of data and for sharing of resources should follow three main principles that is slice isolation, elasticity and end-to-end customization (Afolabi et al., 2018). These same principles can be applied to obtain results, while creating slice for RAN and core networks, which later can provide end-to-end service delivery with isolation.

a) Slice Isolation

Isolation is one the most important properties that should guarantee the performance of a network slice. Isolation will ensure that the tenant's slices are independent, and their traffic will not interfere with the traffic of other tenant slices. Furthermore, isolation will not allow the other tenants to access or use the resources of other tenant's slices in a virtual private network. With Network Function Virtualization, the resource allocation can be limited for particular tenants, and according to their requirements (Doll et al., 2017).

b) Elasticity

Elasticity deals with the dynamic nature of a slice. It enables the optimal usage of the services by reallocating the virtual network components with upscaling and downscaling, as well as reprogramming the functionalities of the control and data elements (Afolabi et al., 2018).

c) End-to-End Customization

Customization substantiates the effective use of shared resources from different tenants. It is enabled by leveraging the NFV, SDN and MANO to provide resource allocation in agile way (Afolabi et al., 2018).

2.2. Network Slicing Enablers

As indicated previously, the transport network segment from network slicing would be enabled by combining best practices from cloud and edge computing, as well as software define networking to provide customizable network function virtualization. This approach envisions postulating slices for services, such as eMBB, URLLC and mMTC. In parallel, it will also support an enhanced feature to allow the traffic related to emergency use cases. At the time of writing of this thesis and the peculiar pandemic outbreak, the VPN traffic for enterprise industries are becoming slow due to congestion of traffic and as a result to home-office environment and

accessing necessary services through their private tunnels (Wolff, 2020). This scenario has invoked researchers to reconsider the design of network slicing to be scalable in the case of emergency situations; that would not impact the 5G services by the sudden burst of internet traffic use by vertical industries, as well as enterprise or private entities.

2.2.1. Software Define Networking

Traditional network devices are developed to possess functionalities premeditated by vendors and receive updates of new features and bug fixes if necessary. This approach of developing has a disadvantage in terms of CAPEX and vendor-specific exertion to service providers. Moreover, once the equipment is designed and deployed there is no other way to perform changes for new features. This drawback in traditional network devices has led the researchers to deploy an architecture that can enable the configurable feature. Consequently, SDN provides this configurable feature that separates the control-plane and data-plane, also known as "underlying network" or "network underlay". In addition to ensuring more granular security (May, 2013), it lowers OPEX and reduce CAPEX subsequently.

Due to significant increase in number of users, applications, machines and internet traffic, the total cost of ownership (TCO)[19] is a prominent challenge for service providers. SDN can be used to tackle this challenge by reducing the TCO with decoupling the data-plane and control-plane. This is achieved by establishing APIs for the programmability of the network. Therefore, this grants the flexibility of network control and allows SDN to provide network virtualization for optimizing the resource allocation in a network. In the case of network slice realization, the SDN can help in three technical conducts. Initially, it will aid in scaling the control and data plane independently. In case of gaming or video streams, the scaling of data-plane is required, while in case of augmented reality it needs scaling for the control-plane, which will be one of the requirements of 5G networks. Secondly, SDN will implement improvements in service design velocity and innovation, which will benefit the business model to expand towards third party applications involvement with rapid deployment. Finally, it will enable flexible and efficient virtualization. This will allow the deployment of logical networks (slices) in the "overlay

[19] TCO is a financial data that include all the system/product cost(it can be hardware, software, upgrading, license purchasing as well operation expenses) (Elvidge & Martucci, 2003).

network[20]". All these three factors can attain our goal in transport segment network slice realization (Kazmi et al., 2019).

a) SDN in Datacenter and Service Provider Network

SDN as a methodology has in its scope a centralized plane, relocating the control-plane functionality from traditional devices and thus allowing the network devices to have forwarding function only (Chayapathi, Hassan, & Shah, 2016). In this fashion, with SDN, the network becomes vendor-neutral heterogeneous network, having a controller interacting with different data-planes developed by various vendors. In 5G, since there will be multitudes of forward-plane devices, it is possible to expand the control plane horizontally, known also as "control plane clusters". These clusters may assure the high availability and scalability of the network. Clusters communicate through Border Gateway Protocol (BGP) or Path Computational Element communication Protocol (PCEP) (IETF, 2009), but these control-plane clusters will act as a single centralized plane.

There are two other planes in SDN apart from control and data-plane i.e. application-plane. While implementing the decoupling of control and data-plane, the control-plane will be managed by an application. In this case, an abstraction of network node information and configuration can be implemented from the management-plane, which is responsible for configuration, fault monitoring and resource management of the network node. The network topology and traffic path information from the control-plane can be abstracted and this portion of information applied to engineer traffic, controlled by an SDN controller.

i. SDN in a datacenter

Datacenter growth has been increased tremendously with the advent of innovations such as smart homes, smart cities and online applications, which lead to form a cloud platform constantly available, and that can be accessed by consumer/vertical industries on demand (Kazmi et al., 2019). These requirements resulted in introducing manifold of physical servers, expanding existing or creating new datacenters. Due to increase in virtual servers, many new challenges transpired, such as scalability, lack of virtual local area networks VLANs i.e. 4096 etc. The VLAN space was further reduced when VLANs are used for isolation of services to support

[20] Overlay network is built on physical infrastructure known as virtual network connecting nodes by means of logical/virtual links. Physical infrastructure is known as underlay network. (F.Naranjo & D.Salazar Ch, 2017)

multi-tenancy. Virtual Extensible LAN (VXLAN) protocol was introduced to resolve this constraint, which provides Layer-2 adjacency between the virtual servers by applying an overlay network, i.e. layer-3 network. This creates a VXLAN IDs, which are scaled up to 16 million segments. VXLAN tunnel has endpoints (VTEP) in an overlay network. Moreover, it is Top of the Rack (ToR) switch or hosted virtual switch. There should be an existing integration of these VTEPs associated with the tenant's virtual machines (VMs) (Chayapathi et al., 2016).For that purpose, this can be achieved if there is a visibility of network and in this case SDN plays a crucial role to abstract the information of traffic engineering and program the forwarding-plane that can be i.e. ToR or tenant's VM. Consequently, SDN will allow communication to the switch and create VTEP interface that is based on VMs instantiated on a server that serves the VMs. This concept can help in abstracting the overlay network tunnel VTEP, which may need to be scalable or deleted by the SDN controller. Figure 41 is an illustration of SDN controller, provisioning VXLAN Network that can be implemented on both physical and virtual machines (Chayapathi et al., 2016; Kazmi et al., 2019).

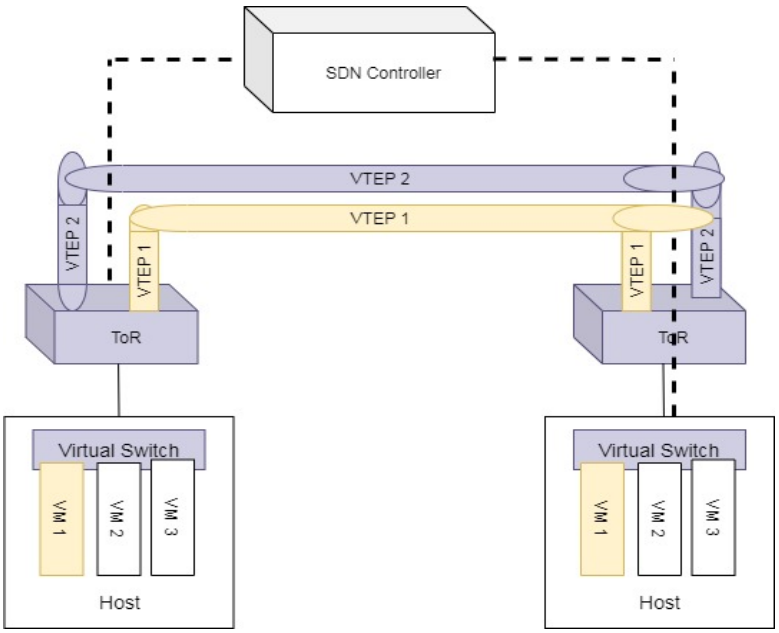

Figure 41. SDN controller provisioning VXLAN (Chayapathi et al., 2016)

ii. SDN in Service provider cloud and security challenges

Service provider (SP) networks are classified into two categories: *edge nodes or provider edge (PE)* and *provider devices (P).* PE routers have features for categorization based on QoS, access

control, routing, and failure detection. In our transport segment slice, it may be necessary to keep information as near to the customer's edge to avoid congestion of private tunnels for repetitive services used by the tenants. The underlay network carries a considerable amount of routing information and address resolution protocol (ARP) caches, which can be used as an abstraction of network topology in a service provider network. An aggregation takes place to P-routers using a high bandwidth links. These links already carry high volumes of data in 4G networks (Chayapathi et al., 2016). In future however, the 5G use-case requirements may disrupt these high bandwidth links as well as the traffic between the tenants' VPNs. To avoid such failures, either physical redundancy is provided, which is substantially expensive solution or as service providers use traffic engineering and traffic steering techniques to steer traffic based on cost and network state. These can optimize the performance of the links for efficient performance. Protocols such as MPLS traffic engineering (MPLS-TE) is commonly used to achieve the goal but there is an overhead of CPU utilization and extensive use of memory, because the devices in the underlay network continuously require end-to-end coordination in the distribution environment (Kazmi et al., 2019).

Nevertheless, the SDN controller can handle such challenges. It has a capacity to abstract the entire network topology link-states, and keep record of bandwidth allocation, as well as implement the decision of traffic engineering. The controller will reduce the overhead of memory usage and CPU utilization of the overall underlay network infrastructure. In service provider cloud environments, enhanced features are introduced such as parameters of QoS and security. If there is any change in parameters, the controller can append all changes to the entire SP underlay edge nodes consistently. In case of Distributed Denial-of-Service attacks (DDoS), the SDN controller can steer the attack away from the standard routing path or virtual private network in 5G, to a path that is not in use, to protect the SP infrastructure. This same method can be implemented by the SDN to control the transport segment of the slice (VPN), for avoiding potential security threats. In case of emergency, it can both divert traffic to other slices with less traffic and isolate the traffic within the VPN with VXLAN or the IP-in-IP protocol concept, as discussed previously in the section about SDN in datacenters (Chayapathi et al., 2016; Kazmi et al., 2019).

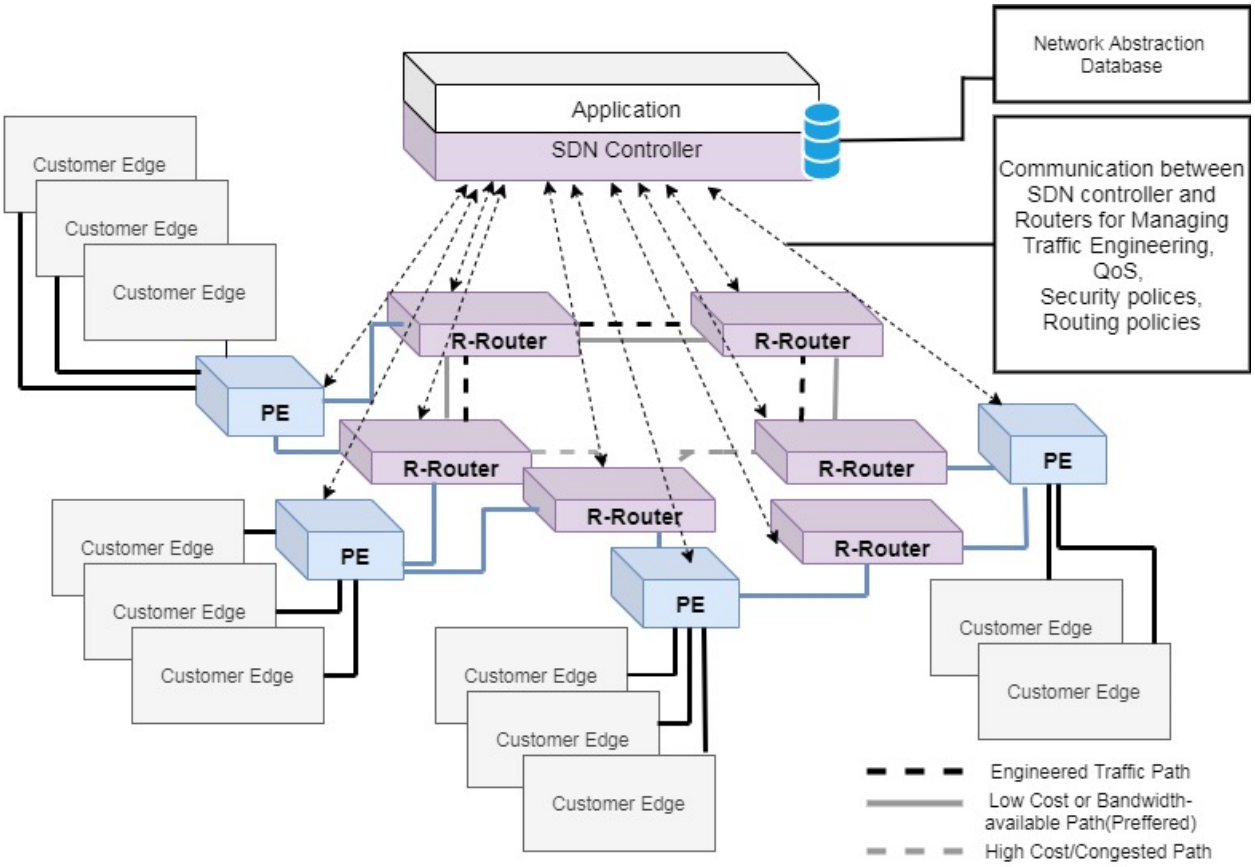

Figure 42. SDN in Service Provider Network (Chayapathi et al., 2016)

a) SDN and the ACTN Framework

Abstraction and Control of TE networks (ACTN) is a framework (Ceccarelli & Lee, 2018), which describes the group of virtual network operations that are orchestrated, while controlling and managing a large scale of traffic engineered (TE) networks. This leverages the process of network programmability, automation, resource allocation and sharing, virtualization and providing end-to-end virtual service connectivity. ACTN project by Internet Engineering Task Force (IETF) is a management architecture to supervise the transport networks (Adrian Farrel, 2018). The main goal is to provide vertical industries and users of telecom services to request and manage the virtual networks built on an underlay network of the service provider. The operations of ACTN are as follows (Onos Projects, 2018):

- Abstraction of underlying network resources to higher layer applications and customers. This information of network resources allows control of the particular resources. In other words, applications and customers of the network can dynamically control the virtual

network, which includes creating, modifying, monitoring, and deleting of virtual networks (VN).

- Orchestration of end-to-end virtual network services, by allowing network resource allocation according to requirements of services by vertical industries and customers.
- Adopting the customer requirements by performing necessary mapping, isolation and enforcing policies in a virtual network abstracted from the underlying network information and overlay network.

Furthermore, this transport network is expanded to include all Traffic Engineering (TE) networks (Ceccarelli & Lee, 2018). A TE network slice from the ACTN reference point of view is a collection of resources to deploy dedicated virtual network over the underlay TE networks.

TE network slicing allows dedicated virtual networks over a common physical infrastructure (Ceccarelli & Lee, 2018). The dedicated resources can be physical or virtual, which are shared among the instances of network slices. These instances are an end-to-end realization of a network slicing. In an ACTN framework, a virtual network (VN) provided by service provider (SP) can be used by customers according to their needs, as it appears for a customer to be a physical network. There can be two ways to abstract the information from VN. The first is to retrieve the information of edge-to-edge links, in which links are referred to these specific VNs and forms a tunnel (Onos Projects, 2018). The tunnels can then be deployed in underlying network either by abstraction of paths and include all inter-domain links, as well as inter-domain paths between the edge to form a transport slice or a TE network slice between the two edge nodes of a customer. Secondly, another approach a service provider can implement is virtual network embedding process, in which abstract topology i.e. abstraction of nodes and links from underlay networks can be used by higher layer networks. The higher-level networks use VN abstraction that includes in this case physical end points, border nodes, and internal nodes as well abstracted nodes. Abstraction topology concept should be implied to enable enhanced VPN services in a 5G network. Figure 43 is an illustration of a simple VPN connectivity in which enterprise offices are interconnected through tunnels represented in red lines, as if it is connected in a direct physical manner, also known as LAN network. This point-to-point connectivity quickly develops a service known as virtual LAN, but once again, the core network is shared, and this infrastructure is not visible to the customer (Onos Projects, 2018).

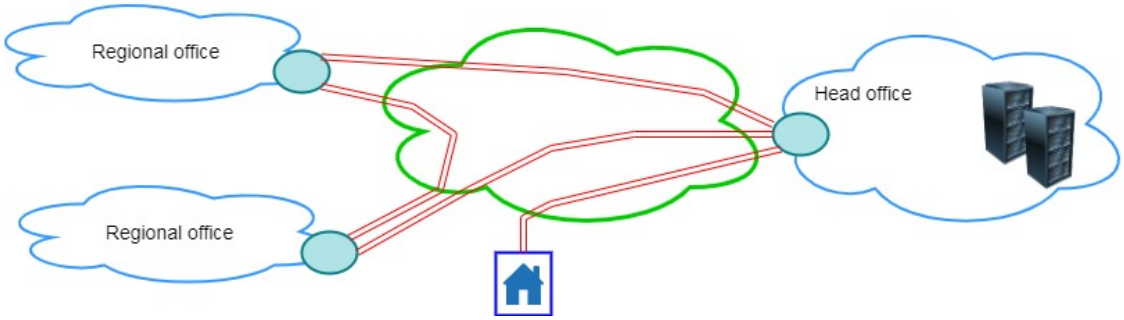

Figure 43. Traditional Virtual Private Network (Adrian Farrel, 2017)

The definitive goal would be to run a private network as a VPN, resembling a real network, privately operated without any notion of underlay networks. Nevertheless, service provider network presence is always required that provides optical nodes or transport nodes. Even in the concept of network slicing in 5G a customer can add or modify the resources of a service as requirements emerge (A Farrel et al., 2016). For this reason, service provider has put a intermediary layer, known as abstraction layer, which gives the information of potential connectivity to provide end to end services. For example, in a case of eMTC, customers may desire to create a network slice for their services, and they need two endpoints for reliable communication. The service provider will then provide an abstract topology to the customer for enabling connectivity from this node to achieve high throughput and reliability but will not provide entire information and details about the underlay network. As shown in Figure 44, topology aggregation is separated into three layers: Client Network, Abstraction Layer Network and Server Network. In this illustration by IETF, client has resources C1, C2, C3 and C4. They want a connectivity from C2 to C3 using some resources from their sites C1 and C4. Clients will use CN1,CN3 and CN5 nodes to build its network, which is a topology provided by service provider, without giving them information of CN2 and CN4, that is a underlay network (A Farrel et al., 2016).

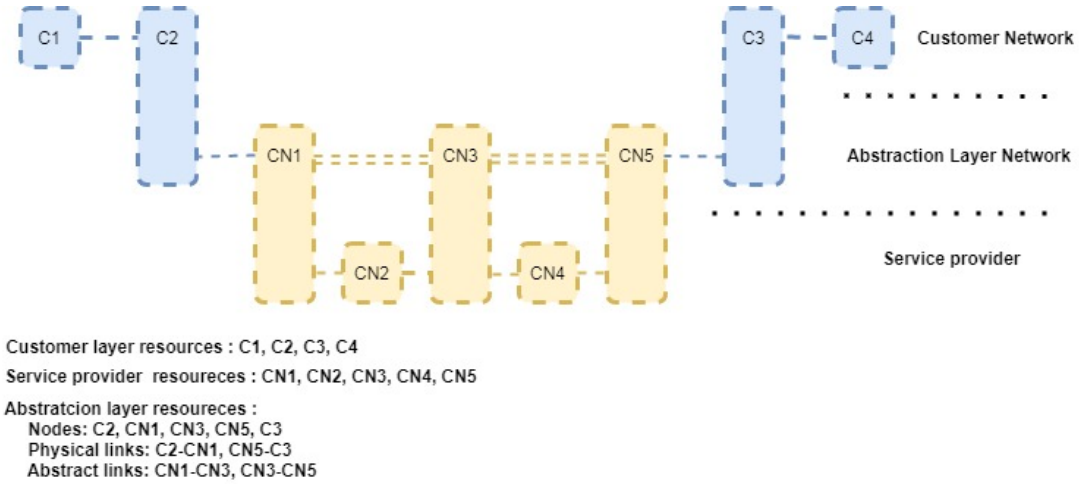

Figure 44. Abstraction topology (IETF, 2016)

Moving forward from abstraction layer, ensures the virtualization of networks (IETF, 2016). This allows for isolation of virtual networks, in which customer uses individual resources and protocols and entertains a control over the same. On the other hand, the service provider may not only support one customer but multiple simultaneously. For this reason, they build multiple abstraction topologies and prompt the customers to build their own virtual private network(s). This virtual network is comprised of physical nodes, physical links, as well as virtual nodes and virtual links. The network then appears to be a special network. In a 5G network, abstraction of networks can be performed with policies provided by Policy Control Function (PCF). This function contains the information of the customer A with eMTC service that needs a high bandwidth along with QoS. Furthermore, customer of Network 1 (see Figure 45) does not want any other networks or vertical industry customers to exploit their resources i.e. complete isolation. The core of 5G, with other functions as NSSF, AMF and SMF, will provide the transport segment slice based on these slices and will form a path of user plane UP through UPF.

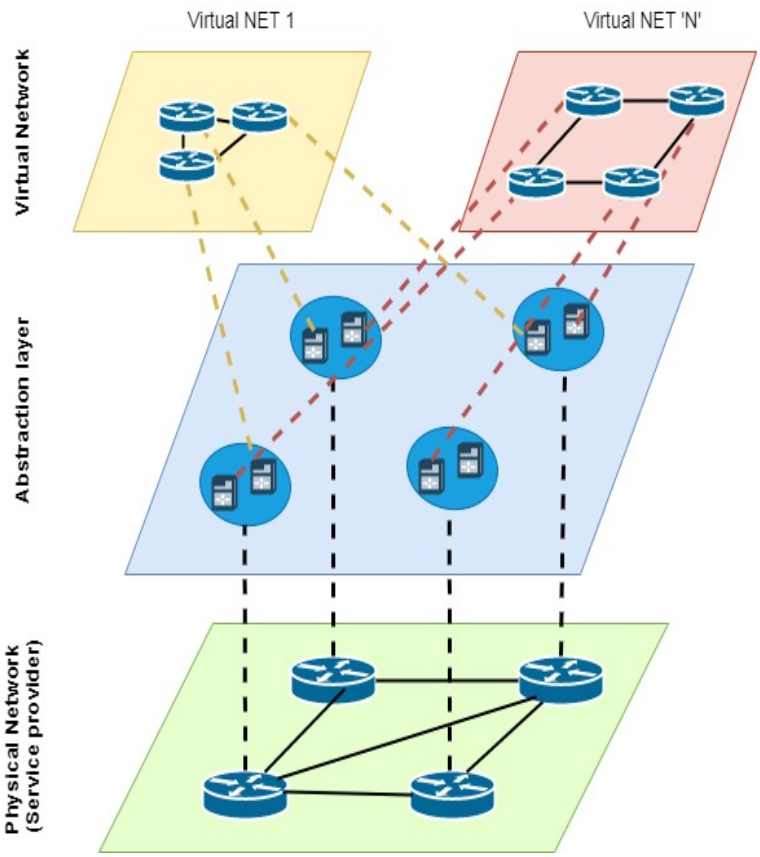

Figure 45. Abstraction leading to Virtualization (courtesy of Metro-haul) (Adrian Farrel, 2019b)

Figure 44 is an illustration of virtual networks, formed through an abstraction layer. This scenario is also known as Traffic-Engineered TE networks (A Farrel et al., 2016). Virtual NET1 and Virtual NET 'N' are from the abstraction layer. Notably on the underlay and overlay networks in Figure 46, the bottom network consists of physical and virtual nodes from the service provider network. The blue and green nodes are physical, and the red ones are virtual nodes connected to each other. On top of this network, tunnels are created, and virtual links highlighted with red paths in the fashion of an overlay network that connects few nodes, but not all the nodes are visible. These tunnels appear as a real physical network to the end customer. Therefore, it maps up into customer network nodes (blue), linked together with red paths that are virtual links connected to virtual nodes (red). In this case, the applications such as VPNs can provide connectivity services. It is feasible to make additional enhanced VPN services with new attributes i.e. super-high bandwidth between nodes, management connectivity by adding, modifying, deleting the nodes as desired.

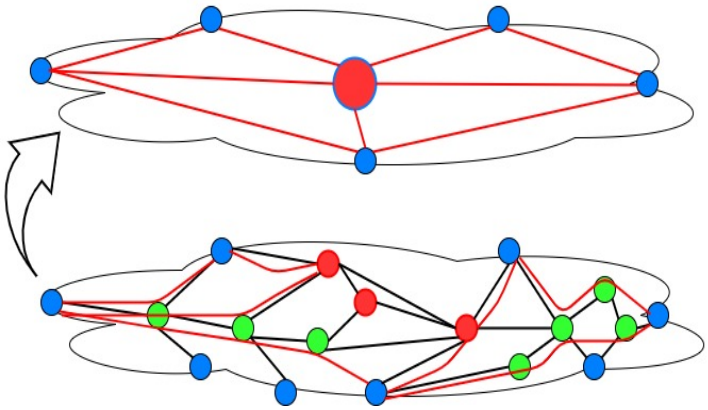

Figure 46. In-depth view of an overlay and underlay network in a transport slice

These concepts define a Network slice, which attempts to partition the resources for enabling a guarantee of QoS and higher bandwidth, either for specific use-case or in case of emergency. When network slicing is discussed, all the industrial partners such as Ericsson, Nokia and academic researchers define the slice from radio network access RAN perspective i.e. 5G radio network slice (Ericsson, 2020; Hirayama, Tsukamoto, Nanba, & Nishimura, 2019). It has been discussed (Barakabitze Alcardo, Ahmad, Mijumbi, & Hines, 2020) to deploy RAN slice from the aspect of Multi-Operator RAN (MORAN) and core slice from Multi-Operator Core Network (MOCN). A spectrum parameter is selected to utilize the resources in deploying RAN slice. A "Cellslice" architecture proposed by (Kokku, Mahindra, Zhang, & Rangarajan, 2013) emphasizes the RAN slice resource virtualization. (Aijaz, 2018) provides the mechanism of Hap-SliceR radio resource framework for resource utilization.

These are some examples in deploying RAN slice, while given less attention to the Transport slice segment in a 5G network. It is essential to provide the same QoS and isolation by RAN slice, and this has led to peering into requirements of underlying networks, i.e. underlay networks that will transport the services either from RAN to other domain network through UPF, or to access functions from the core network of 5G. Thereafter, while slicing the transport network (IP or lower layer), it is not feasible to slice with the same granularity as compared to Radio Access Network (RAN) slicing. The reason for this is that with millions of devices and handsets, the physical and virtual infrastructure of the service provider will consume maximum resources in terms of CPU utilization and storage. A transport slice guarantees the QoS parameters such as bandwidth etc. Beyond this, it will have dedicated or reserved resources in terms of bandwidth,

compute, storage and service functions in an emergency use case scenario. The transport network slice can be built by combining the following three processes:

- Software Define Networking
- Abstraction and Control of TE Networks (ACTN)
- Enhanced VPNs

As described in the SDN concept for Datacenters and service provider networks, the SDN controller is used as a distributed approach (Chayapathi et al., 2016). For transport network slice, an implementation of a distributed approach of SDN controllers is feasible, but for academic research purpose it is not substantially practical. To conclude a specific approach in enabling the slice, we can rather implement a hybrid method, in which a PCE can act as a controller to import the information of the underlay network through BGP-LS (IETF, 2009). The SDN controller allows interaction of network nodes via southbound interfaces to present an abstract topology through RestCONF. BGP is a core routing protocol that makes the entire internet work and it offers a mechanism by which link-state and traffic engineering TE information can be collected and shared to external components of the network. This forms a Traffic Engineering Database (TED) as illustrated in Figure 47.

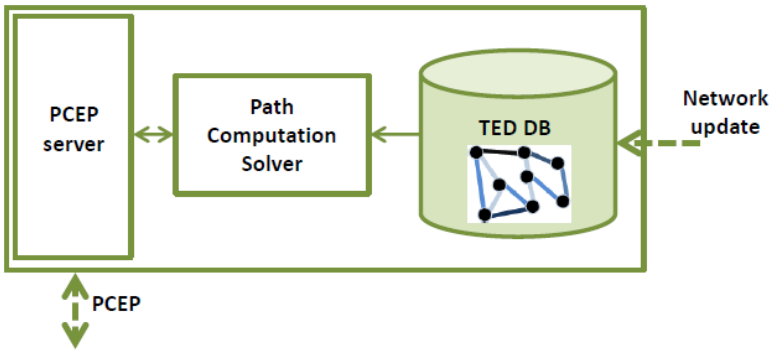

Figure 47. TED database (courtesy of Adrian Farrel) (Adrian Farrel, 2019b)

Furthermore, a Path Computation Element Protocol (PCEP) is used as a standardized protocol, whose job is to delicate the optimal path computation from network elements to an external application. Exclusively, PCEP instantiate paths between the network elements based on TED (IETF, 2009). The BGP and PCEP sessions are established by configuring the transport network. SDN operates on the nodes running MPLS, i.e., interior gateway protocols (IGPs). These nodes have storage link information and once the session is established, the SDN controller can

ascertain link information. Additionally, it can inject the path manually, which provides a hybrid SDN controller mechanism. On these devices, a Path Computation Client (PCC) can be run to instruct routers to program paths or tunnels once PCC-PCE connection is established. In this way, the user can add, update and delete tunnels using RestCONF (IETF, 2009).

Abstraction and Control of TE Network framework works with the SDN-based architecture, detaching the service and network control from the underlying data plane. It consists of three controller types with three different interfaces for communication with corresponding controllers (Adrian Farrel, 2018):

- CNC - Customer Network Controller.
- MDSC - Multi-Domain Service Coordinator.
- PNC - Provisioning network controller.

The interfaces are (Adrian Farrel, 2018):

- CMI - CNC-MDSC interface
- MPI - MDSC-PNC interface
- SBI - South bound interface

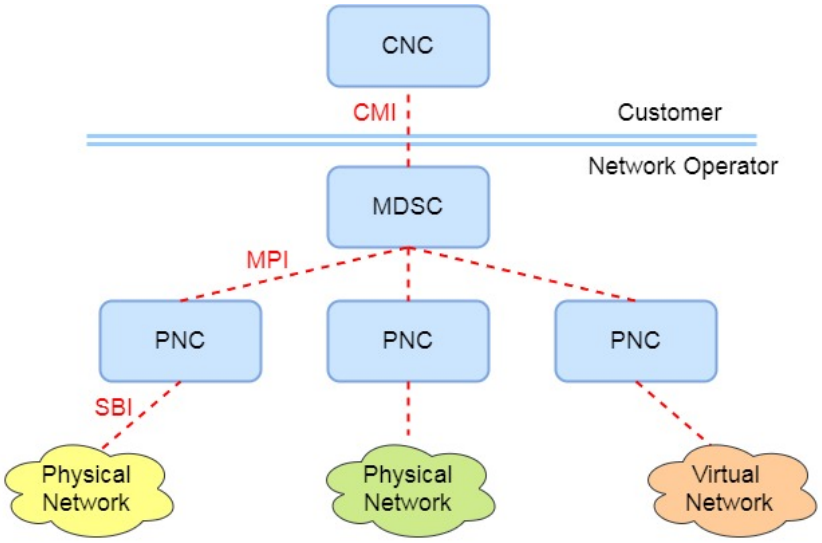

Figure 48. ACTN Architecture Components

i. Customer Network Controller

This is a software-based component, controlled by customer. From here, the customer or the vertical industry user can request to the Network operator for a virtual network through CNC-MDSC interface i.e. CMI. Network operators are responsible for infrastructure that implements provisioning of the network resources and provide it to customers. In this ACTN scenario of network resource provisioning, the service providers are considered as customers that further give services to end users, with the difference that the scenario follows an ACTN framework for VNS (virtual service networks) (Ceccarelli & Lee, 2018). CNC has an information of endpoints, service policy as well as QoS considerations.

ii. Multi-Domain Service Controller

The orchestration of network resources is performed in MDSC. There are two functions in this controller. Initially, it is network functions in which multi-domain coordination and abstraction/virtualization are considered. The second function is a service function, which relates to customer mapping/translation and virtual service coordination. MDSC is the core of an ACTN framework, which provides a communication between CNC and PNC. In this way, it detaches the network and service control from the physical infrastructure of the Network operator to make the network flexible and scalable according to the customer's requirements. Thus, service providers can have multiple MDSCs to provide end-to-end services, connected through multiple PNCs over a common infrastructure of the same network operator (Adrian Farrel, 2018).

iii. Provisioning Network Controller

The third component of ACTN framework is PNC that takes the functionalities provided by the customer from the MDSC through a Southbound Interface (SBI), which instructs the setup of the required connections between the nodes in an underlying network.

According to the previously explained framework, transport slicing is enabled without knowing the underlying network of operators. This has to indicate two endpoints with parameters such as bandwidth, protection, latency etc., to be fully operational for customer end-to-end services (Adrian Farrel, 2018).

3. Implementation

A network slice is a single or a group of logical networks that is comprised of managed resources and network functions (3GPP-TS 28.530, 2019). A logical network provides end-to end services. In 4G, there are Radio Access Network (RAN), Transport Network and core EPC components. Analogously, 5G networks have also three parts of deployment i.e. Radio Access Network (RAN), Transport and core. It is possible to pertain an end to end service in 4G, but with a caveat that the infrastructure is shared. In 5G, to guarantee the end-to-end services that have specific security requirements, the transport layers play an important role that need to be sliced according to the services. It is not straightforward for a customer to understand the complexity of the underlying infrastructure of a network operator. For that purpose, a Differentiated Services Code Point (DSCP) is used as an application that provides mapping but does not consider the essential parameters of isolation, replication, and granularity. For example, if a user has set parameters (QoS, latency, bandwidth) for the service, the 3GPP control-plane selects a slice, 5GI (QoS parameters, such as admission threshold, scheduling weights, link layer protocol configuration, QCI in LTE) and programs the radio access node (gNodeB). The User-plane function (UPF) on the other hand, will not consider the mapping of transport networks (Gsma, 2018). Therein the need for mapping the slice in the transport network segment between the RAN and the Core Network.

To experiment this, we utilize a FlexRAN controller (Mosaic-5G, 2020) to institute two network slices and designate distinct parameters for each. The transport network in the underlay is established on top of OpenStack cloud and integrated with the Neutron service using the Kuryr plugin (OpenStack, 2020a) for running container virtualization as part of the Open vSwitch domain in Neutron (see Figure 49) (OpenStack, 2020b). The containers then communicate with the underlay network directly as if they belong to the bare-metal cluster in the datacenter, without any added overhead (Bruno Dzogovic et al., 2019). As indicated in the figure, the containers are running the Core Network, Baseband Unit and Remote Radio unit as functions of the OpenAirInterface5G, providing end-to-end connectivity between User Equipment and the Internet.

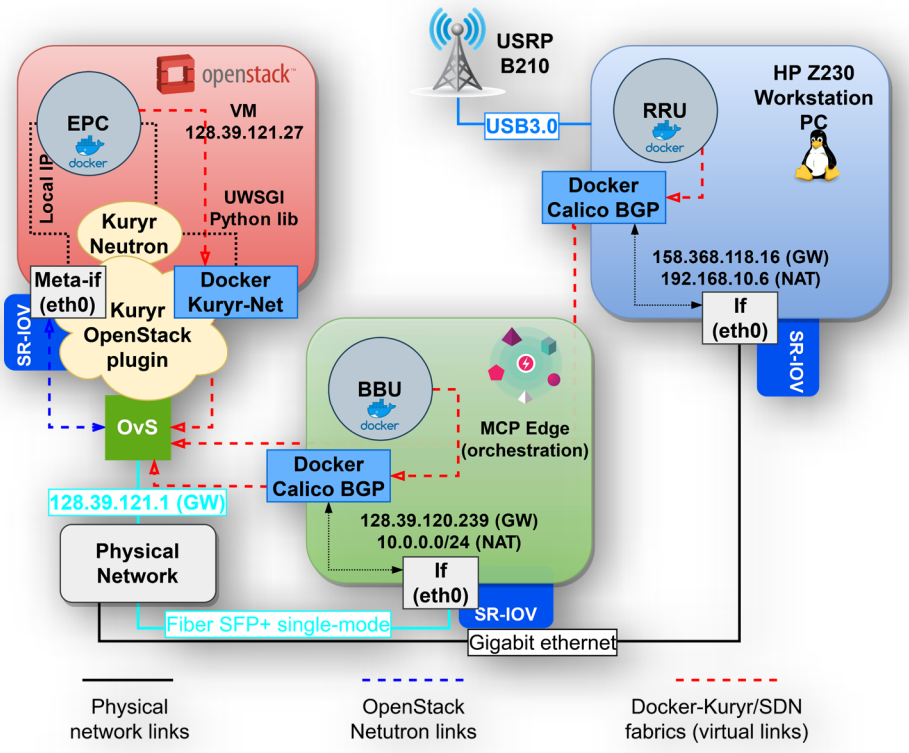

Figure 49. Provisioning a Core Network, Baseband Unit (Centralized Unit) and a Remote Radio Unit (Distributed Unit) in OpenStack cloud. The transport network between the Core Network and the Baseband Unit is integrated using Kuryr plugin for Docker networking (Bruno Dzogovic et al., 2019)

Before commencing with the realization, it is necessary to describe the employment of a transport network slice and its requirements according to the ACTN framework.

3.1. Realization of Encrypted Transport Network Slice

As indicated previously, the ACTN framework defines the ability for customers to deploy private networks without the understanding of the transport network itself. Figure 50 (defines the Virtual network slicing service model, while providing the endpoints with requirements of bandwidth, latency, load-balancing and protection. This same model can play an important role in instituting 5G transport networks by utilizing the CNC, MDSC and PNC controllers.

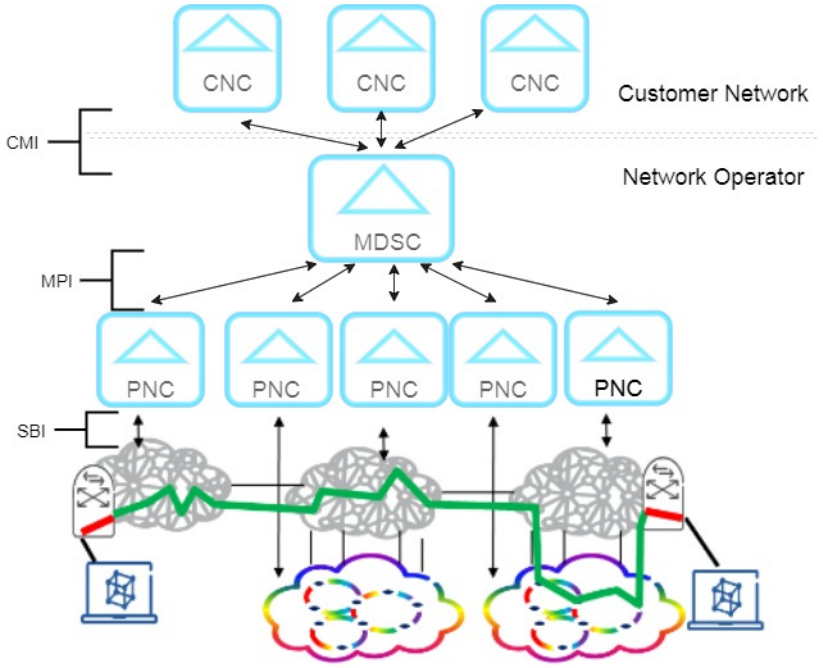

Figure 50. ACTN architecture solution for End-to-End Connection (courtesy Metro-hual) (Adrian Farrel, 2018)

The CMI (CNC-MDSC interface) can be used to interact with a 5G 3GPP mobile network (see section 1.4.2) as well as transport networks. CNC is responsible for 5G 3GPP access-network communication with the underlying network of the 5G infrastructure. CNC can be named as Traffic Provisioning Manager (TPM) (Y. Lee & Kaippallimalil, 2019). With large range of services, high bandwidth, low-latency, and mobility in 5G, the transport networks requirements would change dynamically. In an emergency use case, there is a timespan of maximum few minutes to meet such demands and provide customer an end-to-end connection. For this reason, the backhaul networks should be capable in dynamic reprogramming according to users' requirements. In transport domain, we can have variety of forwarding technologies (IP, MPLS, Segment route etc..). It is thus necessary for an end-to-end transport network to ensure the service quality for these domains. The TPM functions as a CNC from ACTN reference point of view and can be deployed in a carrier network as shown in Figure 51 TPM can be deployed in Mobile Network A - Domain 1 and Domain 2, while the CMIs interfaces are connected to SDN controllers (in this case we call PNC from ACTN framework reference). (Y. Lee & Kaippallimalil, 2019) has further described in IETF document that transport network for 5G has three segments.

- N3 segment between Next generation NodeB (see section 1.4.3) for more details about gNB) and User plane function (UPF),
- N9 mobile connection transport between the two operators, and transport segment that is a backhaul between the two domains,
- N6, a transport segment between the end user/server, which can be a datacenter or head office of an enterprise. Therefore, the transport network can reside between the gNB and UPF, since the data-plane is separated from the control-plane, as well as between two User Plane Functions of another network domain.

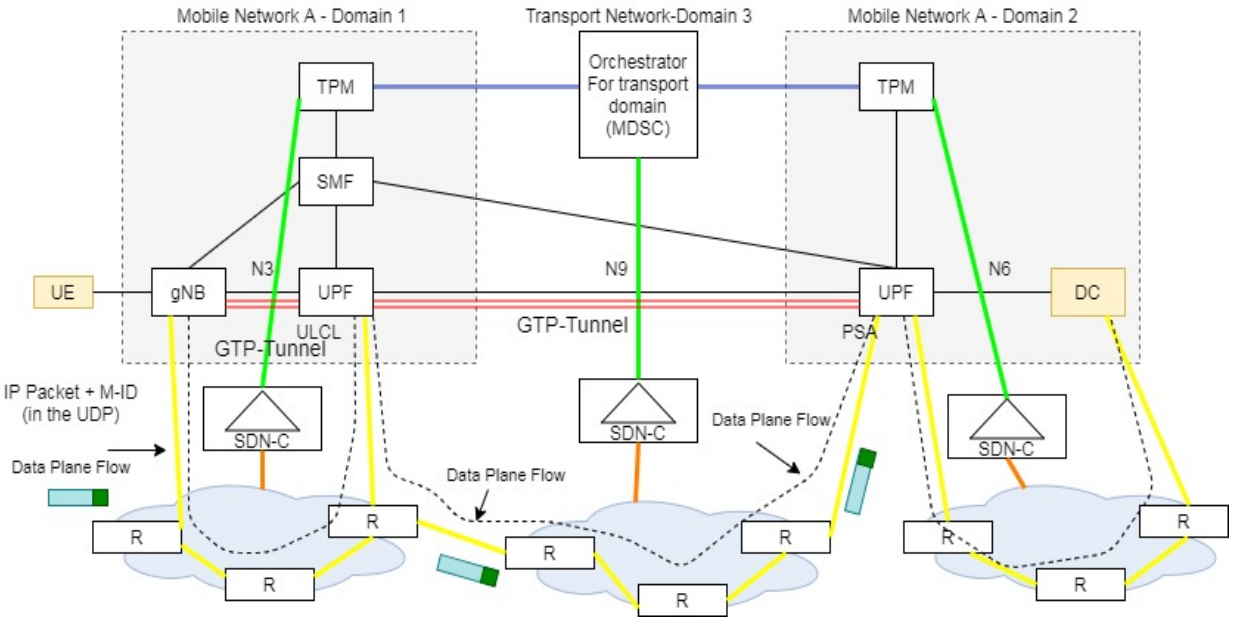

Figure 51. Enhanced 5G Transport Network Architecture (courtesy Y.Lee) (Y. Lee & Kaippallimalil, 2019)

The allocation of resources from the underlying network to customer VPNs shall be based on policies that take into consideration a traffic matrix and QoS, as well as class of network slices with their requirements. The ACTN framework cultivates a policy for deploying 5G transport slices, known as Multi Transport Network Context (MTNC). Consequently, the TPM as a CNC will initiate end-to-end network slice policies based on the identification of MTNC. Due to the MTNC requirements, which include isolation, security and resilience, the controller will provision transport networks for the user traffic between the two endpoints. This MTNC identifier is generated by TPM. As shown in Figure 51, the SDN-C software-defined networking controller will condense the abstraction topology and form a TED. The TED database will have

an access between the TPM and SDN-C. In that case, the TPM coordinates with MDSC to request the virtual networks VNs, which maps the slice later. In this scenario, the orchestrator for Transport Network Slices will be MDSC. When the User connects to the network, TPM will check both the NSSI ID and the traffic matrix. Based on the NSSI (in this case it is MTNC), a slice will be initiated, and an end-to-end path established. Parallel to that, if a customer or user is allowed to access the TPM, they can rearrange the resources, as the abstraction topology can be visible to them since it is abstracted by SDN-C. The end users can then define their own tunnels according to their requirements.

3.2. Network slicing in OpenStack cloud

Thus far, it has been described how SDN controllers operate and enable a network slice with an abstract overview of a network slice itself. This section describes how Network Function Virtualization (NFV) works in context of network slicing. It is a concept of deploying physical nodes into virtual functions that are aimed to provide communication services. For example, in 4G LTE and legacy mobile networks' main components are RAN, Core and transport network equipment provided by the vendors. To deploy the network slices, particularly a transport segment slice, it is necessary to have NFV, which is a prerequisite to network slicing.

NFV makes task faster, efficient, and easier in terms of load-balancing, expanding the network or scaling down, migration, updates without interruption of services. NFV architecture consists of three main functional blocks (Redhat, 2020):

- **Virtualized network functions (VNFs):** VNFs are software applications that handle specific network functions that run in virtual machine (VM) or VMs, containers[21] over the physical infrastructure hardware such as routers, switches.
- **Network functions virtualization infrastructure (NFVi):** this layer involves infrastructure elements (storage, compute resources, router/switches) integrated with hypervisor that runs VNFs.
- **Management, automation and network orchestration (MANO):** manages NFVi and provisions VNFs. Kubernetes is a platform that performs MANO for containers (Intel,

[21] Containers is a software application that has all dependencies (code, libraries) to run the application reliably, isolated from another service, without dependent on the operating system requirement (Linux, Windows) (Docker Inc., 2020b).

2017). In NFV containers provides isolation of services and portability. Docker[22] container technology is one of example that successfully deploy containers in virtualized environment (Intel, 2017).

Figure 52 shows a concept of NFV in 5G. NFV enables elasticity of resources required by virtual networks.

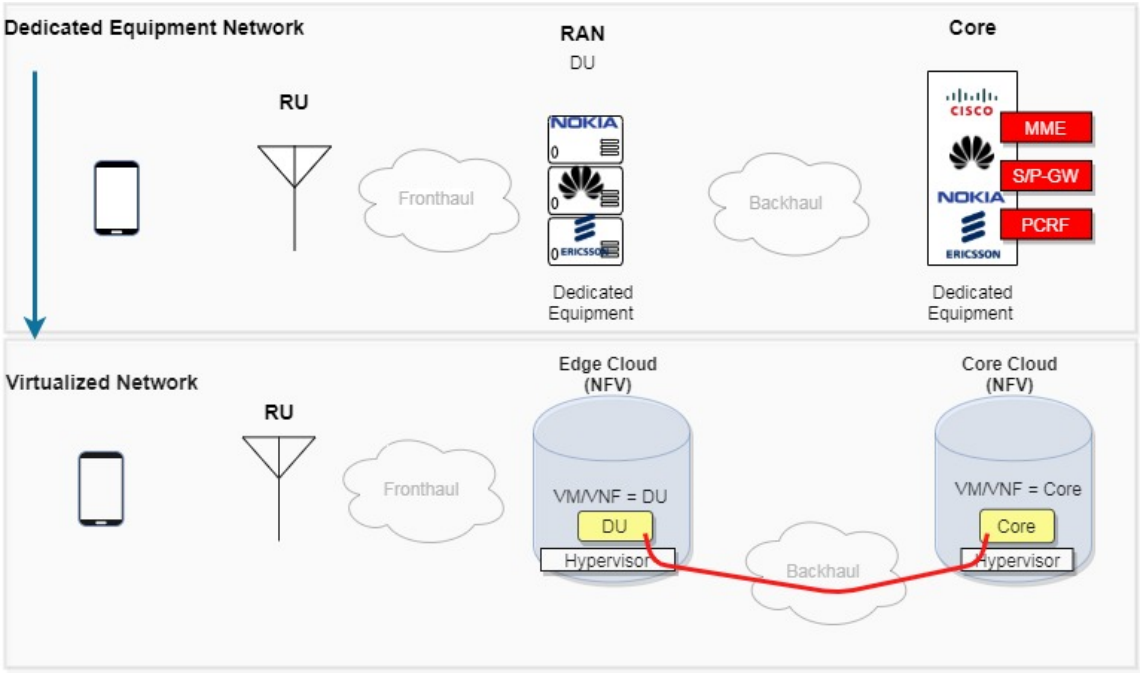

Figure 52. Network with virtualization (Interdynamix Systems, 2018)

Cloud computing on the other hand is very cost-efficient solution model using COTS. It provides a distributed environment for handling large amounts of data and performing large-scale computations. With this paradigm, it is possible to tackle multiple challenges as such isolation, performance, privacy, security and scalability. Before proceeding further, model of clouds is explained and their role in the process of realization of the transport segment of slicing to deploy a tunnel without traffic congestion. Cloud computing provide users with options such as "pay-as-you-go" models, in which a user can utilize hardware equipment such as compute, storage, networking etc. The subscription and scaling based on requirements is typically known as Hardware as a Service (HaaS). The Infrastructure-as-a-Service (IaaS) provides the services in the

[22] Docker is a platform to create, deploy and run applications using containers[21] (Docker Inc., 2020a).

form of computing resources. This type is also incorporating virtualization, which is adequate to provide the resource sharing on a common hardware (multitenancy). If the customer traffic is growing, they can request for more resources and upscale their network. Unlike that, in Platform as a Service (PaaS), there is another layer of computing, which does not require management of the virtual network by the customers, but rather they can deploy their own applications and host services. Service providers in this case are managing the hardware and maintain the network for specific customers. In a 5G network, the functions will be stored in one cloud, the Core Network in another cloud and the RAN functionalities can be deployed in a different cloud. Presumably, these are private clouds and users want to access the services defined in the latter, for example the services of eMMB stored in core cloud. When this function is built, it has some parameters to be able to guarantee the performance, which can be high data rate, reliability, QoS, etc. When customer want to access the services with such parameters, as discussed in the previous section 2.2.1, we can assign MTNC ID to this type of service and TPM (a type of Customer Network Controller in an ACTN framework). This will consequently initiate the underlying resources through MDSC and represent the transport network slice of that user with complete isolation and security. Even if the customer is using the IaaS platform, all this information can be used and defined within the Policy Control Function (PCF) and Network Slice Selection Function(NSSF) to determine the type of service and build the backhaul network, prior to the instantiation of other services being used by the customer. Consequently, to this explain this, four case scenarios are considered, considering the function virtualization and cloudification that lead to the realization of transport segment slice:

- Ultra-High Definition UHD slice: This involves the 5G core and user plane function (UPF) implemented in Edge cloud and cache server and RAN DU, as well as MVO server in the core.
- Massive IoT slice: Does not require high bandwidth. It is a light service slice and thus IoT services are implemented in the cloud core of the operator’s infrastructure.
- Ultra-reliable low latency communication slice: 5G core and associated virtualized functions are deployed in edge cloud to minimize delay.

We are introducing MEC multi-access Edge computing in the slice for performance and efficient access of services by the users or vertical industries. In parallel, there is a separate

MTNC ID defined by the ACTN framework for each slice, which is an isolation of services from different slices. At this point, the network slice is end-to-end, i.e. a RAN slice, Transport slice and core slice; these slices can be managed independently. Moreover, the network functions in the RAN core or transport slice can be managed independently. This may include resources and performance, QoS and other diverse variety of parameters. Although our focus is on the transport segment of the slice, we consider the other part of slices in the RAN and the core, concisely. In the RAN, a slice can be achieved by splitting up channels. For example, it is possible to allocate 100MHz channel spectrum for an eMBB slice, for V2X we decide to put 50MHz channel and since an IoT since does not require high bandwidth, substantial amount of the overall communication can be put in a 20MHz channel. In the core however, because it is implemented on the cloud computing platform, we may devise core network functions (NF) at edge cloud or edge datacenter, or we may have NF in the local datacenter serving the region by providing services locally. It is also possible to acquire network functions in centralized data center, accordingly. Moreover, we allocate both user-plane (UP) functions and core plane functions instances to specific network slice that do not interfere other instances in the slice. Figure 53 is an illustration of NF in NFV environment (Wang, 2015; Zhang, 2018).

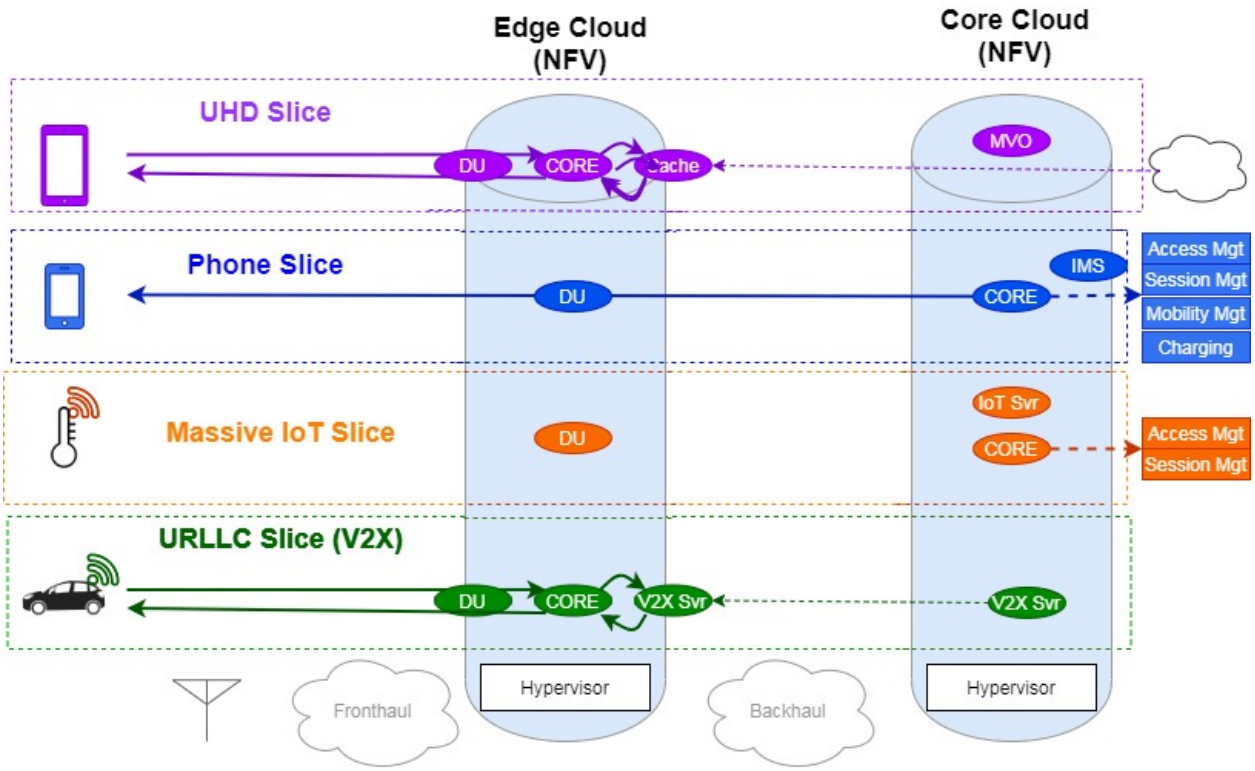

Figure 53. NFV enabling a perquisite for Network slices (courtesy Interdynamix) (Interdynamix Systems, 2018)

Continuing with the previously elucidated use cases, virtualized functions are defined and are being implemented in different locations, i.e. edge and core. We will apply the ACTN framework on these uses cases and observe how the hard isolation can be achieved. For that purpose, an SDN controller is implemented, which in case of soft isolation can support services such as MPLS, VxLAN and Segment Routing (SR) based on VPN and VLAN. In soft isolation, the slices are isolated in such a way that they cannot interfere each other in a same tunnel or a traffic-engineered path, which is the shortest path from edge DC to the core cloud. The resources (CPU or storage) are being shared in this slice isolation scenario.

With the enhanced VPN feature of isolation, the hard isolation has one advantage over the soft isolation, that is to have a complete separation of underlying network so the traffic of particular network slices can use the distinct network resources (Adrian Farrel, 2019a). To rectify this, we resort to the framework provided in the realization of transport network slice that is previously described in section 2.2.1, where a vertical industry customer provides the input of their requirements. Presumably, the UHD slice is given with MTNC ID = 1, the slice for phone access as MTNC ID = 2, Massive IoT slice with MTNC ID = 3 and URLLC slice MTNC ID = 4. These ID numbers will be appended in TPM (CNC controller) and communicated with the MDSC. Since the SDN-C has an abstraction of traffic-engineered network topology, it will assign the path to these slices. This same SDN-C has the logical abstraction of the topology in case the vertical industry does not want hard isolation and can build a tunnel based on MPLS or VxLAN VPN. Nevertheless, since we are focused on hard isolation the vertical industry can choose their private tunnel between the two endpoints, which allows them for a complete protection of their data without concerns regarding the interference of other slices' traffic shall consume their resources or bandwidth. In case if a vertical industry desires an instance of a slice, they can differentiate the slice with a concept of differentiator with introducing an additional parameter of MTNC sub-differentiator i.e. MTNC ID 1.1 (where this case represents another instance of slice MTNC 1).

Similarly, in case of a sudden burst of traffic occurring in a UHD slice, the vertical industry customer has an overview of the abstract topology; based on this, they can add additional resources for smooth transmission of data in their tunnel. Furthermore, this will not affect other vertical industry customers since their traffic path is completely independent from the other slices. Figure 54 is an illustration of transport network slice realization on a SDN/NFV level.

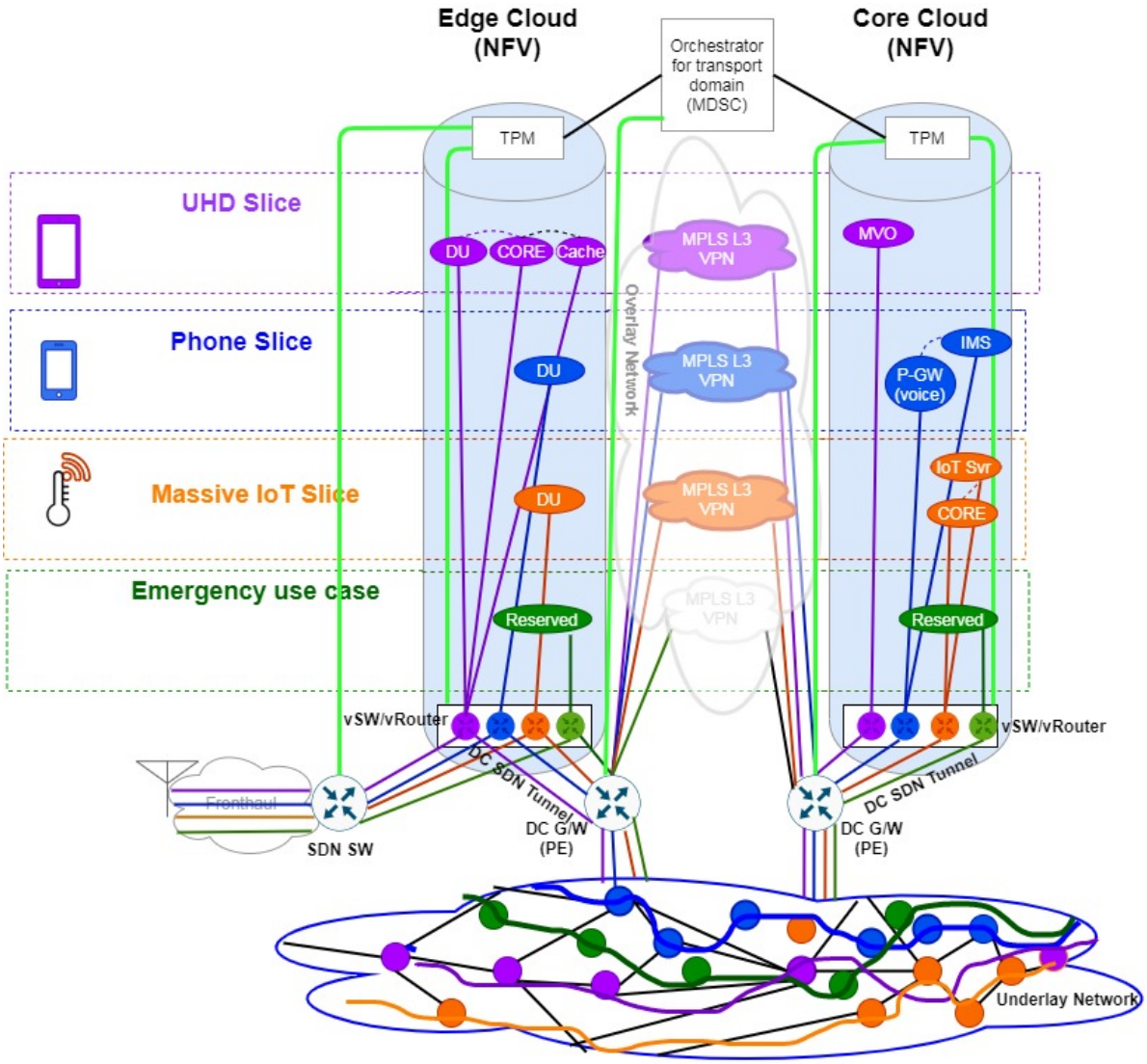

Figure 54. Hard isolation and Soft isolation using SDN/NFV

3.3. Instantiating a VPN tunnel between the Core Network and the Centralized Unit in the cloud

As described previously, we establish a tunnel between the two endpoints in the cloud, which are the Core Network (5GC) and the Centralized Unit (Baseband Unit). For securing the tunnel, AES-256 encryption is used and its impact on the performance measured. The tunnel should be able to correspond to the virtual functions instantiated by the SDN controller. In this case, the FlexRAN controller has the purpose of maintaining connection between multiple MME instances, as well as distinct HSS databases with mobile users in each. The users and the overlay network are not concerned with the transport segment and its functionality, which means that the

users can utilize their own separate VPN applications on their User Equipment. However, the User Equipment is categorized in different slices according to the use-case. In this situation, the initial use-case of a User Equipment is mobile phones communication through the 5G network, whereas others can be IoT devices belonging to different network slice. Each of these will be grouped distinctly in different HSS database and their IMSI code used to assign the same to the specific network slice at the SDN controller (Figure 55).

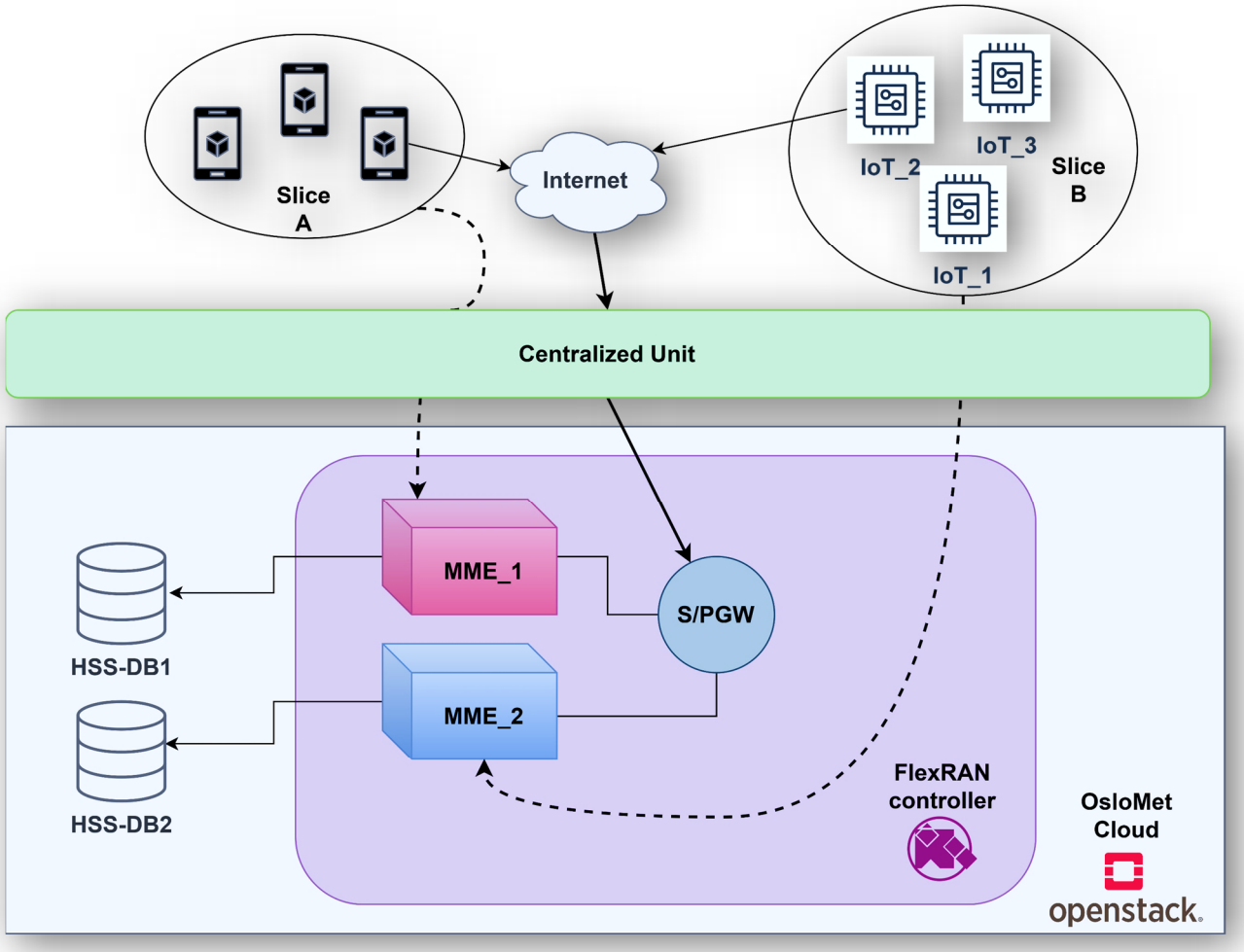

Figure 55. Software-Defined Cloud Radio Access Network topology at the Oslo Metropolitan University

The S/PGW (Service / Packet Gateway) in the mobile core network communicates directly with the Centralized Unit, that is the Baseband Unit for processing functions related to the gNB base station (Radio Resource Control, assigning bearers, link-layer functions etc.). The communication between two distinct virtualized environments (VMs or containers) in one cloud

is usually sufficiently secure by the provider, but the security level is not absolute and therein not absolutely guaranteed in the Service Level Agreement. For that purpose, to strengthen the security between the Core Network and Centralized Unit, it is necessary to establish a specific policy. The policy will dictate the necessary arguments for entities in that network, which will be allowed or disallowed to access specific endpoints, such as the HSS databases. The HSS database is a MySQL database, running with Apache and Phpmyadmin. Securing a database by assigning listening address to 127.0.0.1 may seem sufficient, but this can easily become a disadvantage when there is no specific policy for accessing other locations in the same virtualized environment. To address this, we move the HSS database(s) outside of the virtualized environment into another one and we assign private communication network. This network is filtered by the OpenStack's firewall, with ports listening only in the local network for accessing the database and the FlexRAN controller operates on the level of another Virtual Machine or container. Dissecting the architecture in this manner allows for more granular control over the services in terms of assigning more specific network policies, Role Based Access Control (RBAC), as well as replication in terms of microservices architecture and rolling updates using Continuous Integration / Continuous Delivery pipelines (CI/CD).

In some instances, this model may require encryption between the endpoints where the whole user traffic flows, to disallow specific user (that may be attackers) to penetrate into the core network and compromise another slice being served by the same one. For reference, if we assume that the "Slice B" from Figure 55 is industrial production slice for automation and robotics, malicious users that try to penetrate into that slice from "Slice A" can face additional difficulty to perform the given operation. In a heterogeneous environment, there is a requirement for advanced machine learning anomaly detection systems to substitute or supplement a firewall and prevent such attacks. Nevertheless, when utilizing encryption, the attackers may relinquish the intended actions due to the impossibility of cracking i.e. AES-256 without proper supercomputing power.

To conclude the experimentation, we proceed further with testing the difference between the network slices in containerized environments as represented in figure Figure 56.

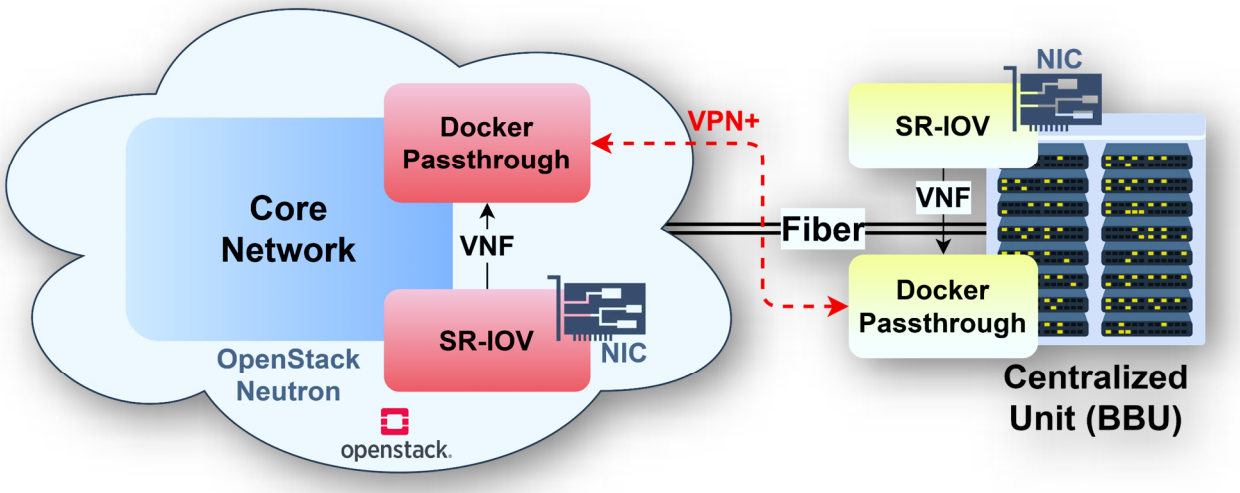

Figure 56. Enhanced VPN on a transport layer between the two Docker instances of the 5G Core Network and the Centralized Unit in OsloMet's datacenter (a cloud and edge server rack)

In this instance, to allow Docker to access the SR-IOV drivers from the Network Interface Card (NIC) virtualization plane, a special driver needs to be incorporated and allowed for docker to use any instance of Virtual Network Functions (VNFs) of the SR-IOV (Mellanox Technologies, 2020). This will allow the Docker instance of the container running the Core Network and the Centralized Unit to access the network underlay, that is Neutron's Open vSwitch in OpenStack cloud. To avoid compatibility issues, we set the Neutron driver for network virtualization to be the Intel's SR-IOV and adjust only one virtual function per instance. This will ensure that we get unobstructed deployment and the implementation of VPN between these instances can be achieved without adding overlay on top of the existing network. As mentioned previously, Docker containers can communicate with the Neutron service in OpenStack using the Kuryr plugin. To allow this, we set the proper ID of the user and VM instance running the Core Network. The container performs authentication through the Kuryr plugin via the OpenStack's Keystone service for handling the authentication procedures in the cloud. In addition to that, as represented in Figure 56, the connection between the physical nodes (the specific compute node running the 5G Core Network and the physical machine running the Centralized Unit, that is a Dell server) is based on a direct fiber link, allowing for uninterrupted 10Gbit/s communication.

Consequently, we measure the network performance using the iPerf3 tool (iPerf3, 2020) to represent a relative figure of the designed solution. These tests are arbitrary, and they do not depict the realistic conditions of a production environment, as the same serve for elucidating the feasibility of its implementation. Additionally, our focus is not on benchmarking the given solution, but rather showing the method of design and realization. For better understanding the impacts an encrypted transport layer network can have on the backhaul connection between the Centralized Unit node and the Core Network node, we perform tests on the direct connection between the nodes and compare these results to the tests obtained from a tunneled link between the two instances using encryption. We expect a substantial degradation in the traffic bandwidth, most likely attributed to the computational requirements for encrypting each packet transmitted using AES-256 encryption.

The radio frontend implementation is achieved by utilizing Software-Defined Radio USRP B210 (Ettus Research, 2020) and connecting UEs (Mobile Phones) to the particular instances of the running base station. As indicated in Figure 57. The Remote Radio Head, which is the Distributed Unit (DU). To achieve slicing, the distributed unit needs to separate the functionality of the base station and move some part into the cloud (the Centralized Unit, which runs in the cloud and connected via SSH through the Baseband Unit – BBU designated machine on the figure) for baseband processing, also called “Option-7” according to 3GPP specifications (ITU, 2018). This allows to minimize the impact on the instantiated Radio Frontend peripherals, as the baseband processing requires immense computational resources to conduct calculations for FFT/IFFT (Fast Fourier Transform) and the radio components. The scope of the 64-QAM constellation transfer and radio bearers are evident from the right screen in Figure 57, whereas the core network running in the cloud is depicted on the left side.

Figure 57. The radio-frontend view and the connection of devices using Software-Defined Radio on Band3 (1800 MHz) and Band7 (2600 MHz)

To achieve a normal transfer and successful further testing, we calibrate a UE SIM card to authenticate into the HSS database and gain connection to the internet. With this, we establish an essential model of a 5G core, Centralized Unit and Distributed Unit for experimenting realistic scenarios. Nevertheless, our focus is not on the radio frontend and the Distributed Unit, but rather the connection between the Cloud Core network and the Centralized Unit in the C-RAN.

Figure 58. A successful DIAMETER authentication procedure between a virtualized MME and HSS instance

To run the core network in the cloud, a pre-installed image is pulled from the Docker repository and instantiated with the following Docker command:

```
docker run -d -it --net=oainet --name=oai_epc --restart=unless-stopped --
expose=1-9000 -h=epc --privileged=true --cap-add=ALL -v /dev:/dev -v
/lib/modules:/lib/modules brunodzogovic/oai_epc
```

This runs a container from the “brunodzogovic” repository and adds it in a network named “oainet”. The “oainet” network is created with the Kuryr plugin in order to obtain an access to the OpenStack’s Neutron service directly and bypass the overlay network in which the container would have run otherwise on top of the particular Virtual Machine instance. With this, one layer of traffic disturbance is eliminated. The core network containers need to have multiple ports opened for initiating GTP tunnels and allowing radio bearers to be assigned for specific UE authenticated in the core. This indicates a security issue and therefore requires better isolation of the environment. In this case, as the command suggests, the ports 1-9000 are exposed from the container to the host machine. Furthermore, we run the Centralized Unit container with the

Docker SR-IOV plugin (Docker, 2020). With this, it is feasible to achieve a certain level of hardware-based isolation of the Virtual Network Functions (VNFs). To instantiate the plugin, we refer to the command:

```
docker run -v /run/docker/plugins:/run/docker/plugins -v
/etc/docker:/etc/docker -v /var/run:/var/run --net=host --privileged
rdma/sriov-plugin
```

For running the container in the cloud and allowing passthrough via the desired virtual function, as described previously in Figure 56, we exploit the parent interface of the virtual machine that maps the virtual function from the Kuryr networking plugin to directly access the Neutron service in OpenStack. Creating the network is as follows:

```
docker network create -d sriov --subnet=192.168.0.0/24 -o netdevice=eno1 -o
vlan=100 CU
```

The specific network will allow containers to connect at the same subnet and communicate as if they are directly connected to the NIC card of the underlying hardware nodes. The powerful aspect of the SR-IOV virtualization is that it achieves near hardware-level performance, without almost any overhead and this translates directly into better performance for the virtualization plane. As reported, the CPU overhead that SR-IOV imposes on the virtualization plane in a VM is approximately 1.76% and without sacrificing network throughput (Y. Dong et al., 2010). According to that fact, the performance of the SR-IOV direct link between the Centralized Unit and the cloudified Core Network will be tested in order to determine the difference between the hardware-level performance of a direct fiber link and containerized environment.

Conclusively, an enhanced VPN deployment is crafted between the two SR-IOV endpoints in the Docker containers, allowing for encrypted communication without a MPLS-BGP encapsulation. This will provide a clear understanding of the impact of CPU-based encryption using the AES-256 algorithm on the performance of the underlying transport network fabrics. As SR-IOV can achieve a hardware-level isolation between endpoints using VLAN segmentation (see Figure 59), this may prove to be insufficient in a multitenant environment, as additional containers and services can then access the 5G core, which should be isolated. One method for maintaining isolation is by policy enforcement, guiding traffic to the 5GC core from only sources that should access it (i.e. a Centralized Unit or multiple Centralized Units). For operators who desire an additional layer of isolation, despite the underlying policy, a VPN instances are established between the SR-IOV endpoints.

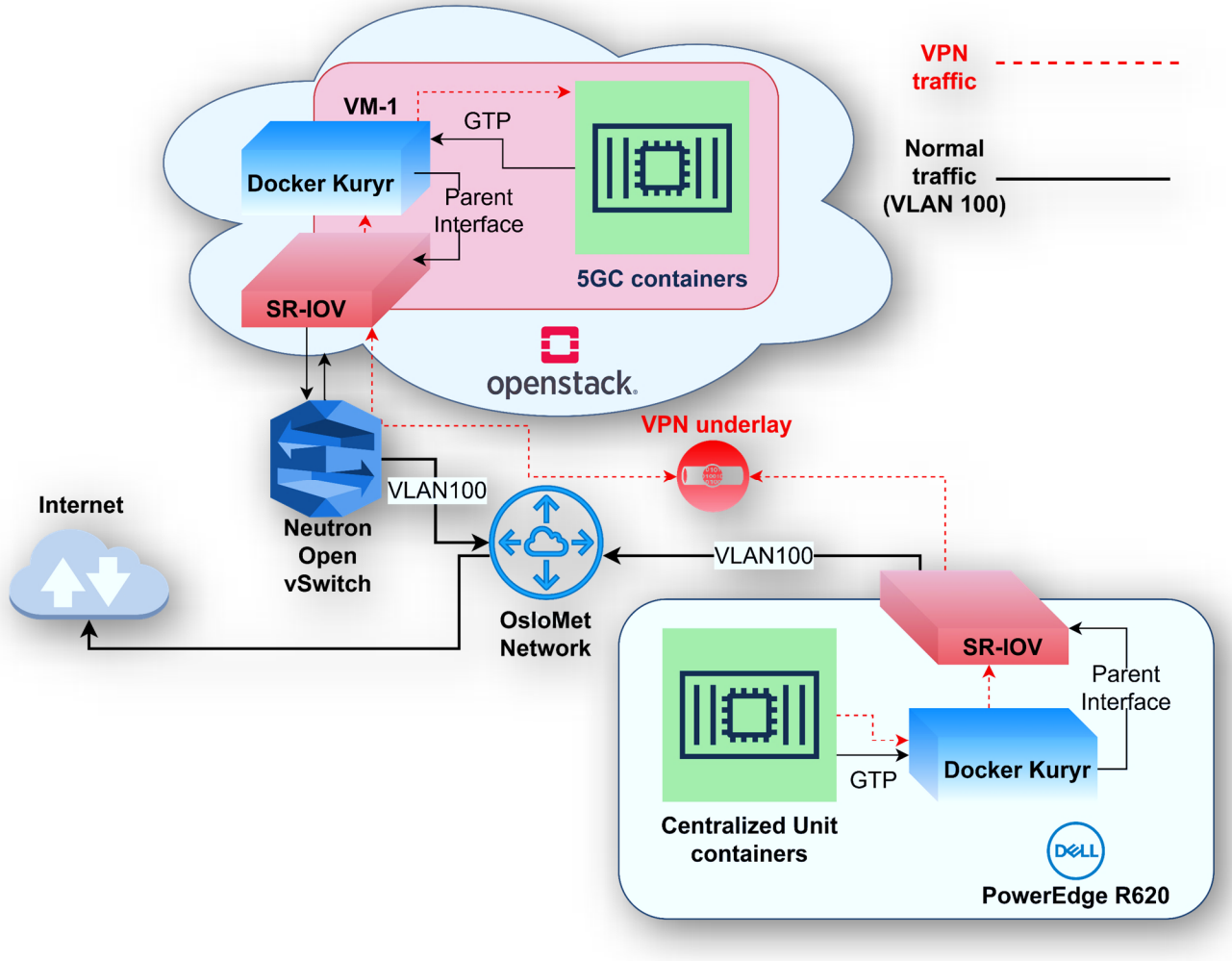

Figure 59. VLAN segmentation using SR-IOV and VPN instance in the transport network between the Centralized Unit containers and the 5G Core Network in the cloud

4. Evaluation

The evaluation of the deployment is categorized in two instances and two test scenarios in each. The first instance is a test of the underlying infrastructure, to establish an overview of the maximal possible performance at the hardware level. Furthermore, this hardware-level testing instance performs evaluation of a single-stream and multiple parallel streams in the network, to compare the resource utilization that a traffic can generate. With such a notion, we can establish whether the instantiated enhanced transport-level VPN between network slices can induce a hardware-level overhead, rendering the solution peculiar and contemptable of implementation for services that do not require critical security isolation and are strictly performance-depending. Consequently, the second instance are tests on the virtualization layer after instantiating the SR-IOV drivers and running the transport network VPN+ between the Core Network and the Centralized Unit in Docker containers. Last but not least, we conclude with comparing the offset between the overhead induced by the VPN+ on the hardware-level virtualization, traffic performance and CPU resource utilization in parallel, to determine the consequential effects and whether the additional encrypted isolation of network slices is a viable option for various 5G use-cases.

Before conducting the evaluation of the throughput in the network, the Maximum Transmitted Unit (MTU) is adjusted to 9000 from the default 1500 value. This is due the fact that the traffic between the Centralized Unit and the 5G Core Network is encapsulated to support the transmission of GTP traffic in an extended IP header. With a MTU value of 1500, the traffic between the UE (User Equipment) and the Internet will experience packet drop, jitter and additive error, therefore it is adjusted accordingly on all interfaces, including the SR-IOV physical functions, virtual functions, as well as the container-plane network.

The operating systems on which the containers are initialized is Linux 18.04, with low-latency kernel version 14.18.0-25 identical on both sides. The Core Network needs a stable environment to operate and thus any fluctuations incurred by the CPU (Central Processing Unit) on a hardware level in terms of frequency, power draw, heat-induced inconsistency and multitenancy can cause unpredictable results and also varied outcomes in terms of experimental reproducibility. To minimize this likelihood, we disable the Intel CPU hyperthreading and allow a thread-per-core to be initialized for running the Core Network and the Centralized Unit equally. In terms of scaling

the central processing, the CPU is overclocked to 3.0 GHz speed and locked, therefore, no additional fluctuations are expected. This can be reinforced with disabling any power-saving states for putting the CPU in idling mode and disabling the turbo-boost option on a hardware level (in the BIOS and Linux Kernel).

```
echo 1 > /sys/devices/system/cpu/intel_pstate/no_turbo
echo 0 > /sys/devices/system/cpu/cpufreq/boost
cpupower frequency-set -g performance
```

The last line in the command sets a governor of the CPU frequency to run at maximum performance constantly.

4.1. Testing the underlying infrastructure

4.1.1. Single-stream hardware-level tests

In the initial testing iterations for attaining a control reference group, the TCP slow start is not being excluded, as the same accounts for a more realistic perspective of the underlying hardware performance. A scenario like that equals an environment more similar to a production milieu in which 5G networks are instantiated. The iPerf3 is ran with the following arguments:

```
iperf3 -t 300 -i 1 -w 32M -c 10.0.0.1
```

The time of execution is 5 minutes (300 seconds), with an interval of transmission each 1 second. The TCP buffer size is set as a constant to 32 Megabytes and the test runs as a client-server model with the server being executed at the 10.0.0.1 host, which is the Core Network and the 10.0.0.2 is the Centralized Unit host. At the initial stage, a testing of the infrastructure is performed to assess the performance of the bare-metal communication, without any layers of virtualization. These results summarize the maximum possible bandwidth possible over a 10Gbit/s fiber link between two NIC interfaces, as seen in Figure 60.

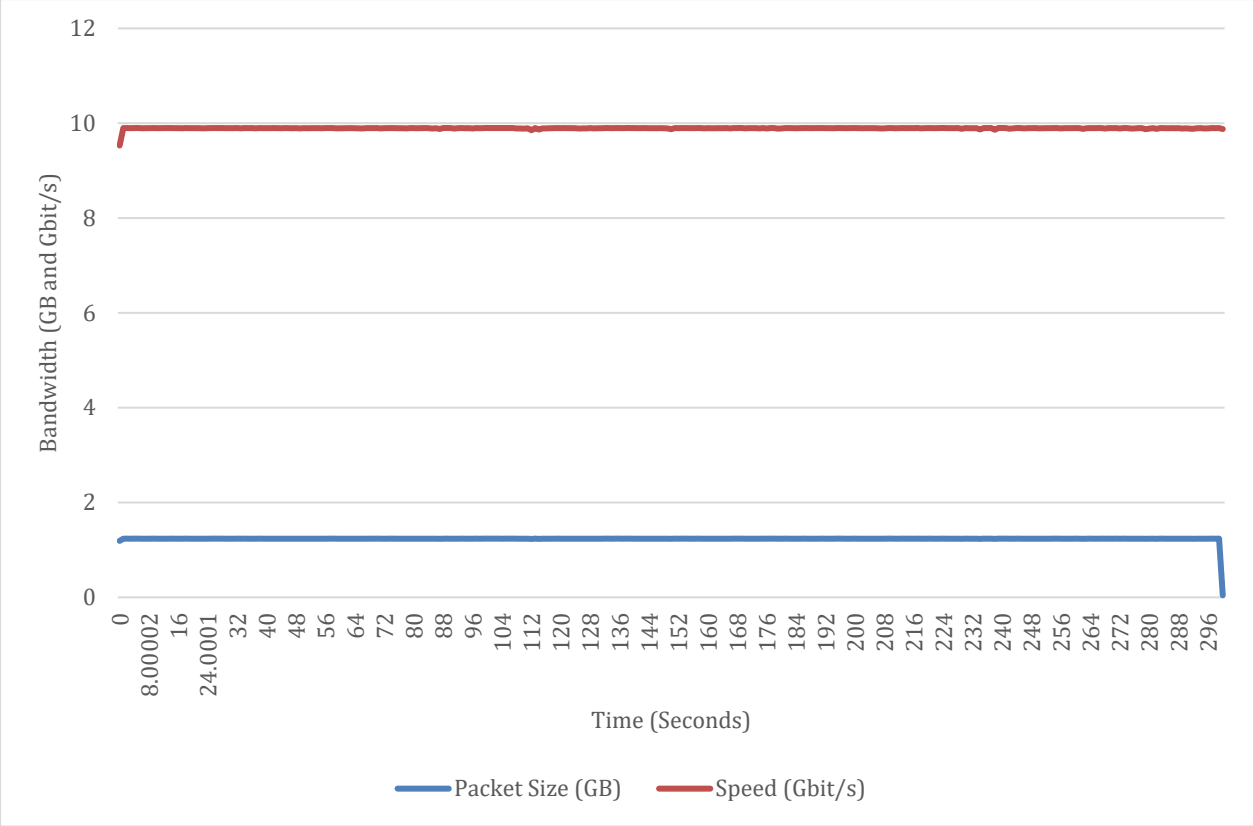

Figure 60. iPerf3 single-stream bandwidth test of 300 seconds on hardware-level infrastructure to evaluate the fiber link-level performance

Table 6 summarizes the test results from a single-stream performance benchmark. The average stream bandwidth achieved by transferring 371 GB of data is 9.89 Gbit/s on the 10Gbps link. This indicates that there is a potential for optimizing the hardware to achieve better results, but that is not the primary goal of this research. The *eno1* interfaces from the physical NICs at both endpoints are supporting jumbo frames of MTU=9000, and this results in a sacrifice of the total bandwidth the physical link can achieve.

Table 6. Summarized single-stream hardware-level performance tests

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (GB)	Average Stream Bandwidth (Gbit/s)	Retransmissions (total)
sender	5	0	300.036	300.036	371.0426782	9.89328	0
receiver	5	0	300.036	300.036	371.0424993	9.89327	0

Additionally, the CPU utilization is measured and summarized in Table 7. The total CPU utilization on the 5GC physical node is 61.9658%, and the current user that is logged in the

session uses 8.23386% of the overall CPU. This amounts for the necessary processes to run in Linux, together with the load induced by the transmission during the time of measuring. The reason why we summarize the overall user CPU utilization is because in reality, the Core Network will be instantiated on top of hardware/virtual environments, which also need a particular operating system in order to run. This is a variable that depends on the operating system, kernel, running processes in the background etc. These processes utilize almost half of the system's CPU capacity, particularly 53.7319% of the total available. It should be also noted that all the other constituents of the 5G core network are sharing these resources, increasing on the CPU overhead in dependence of multiple factors: numbers of UE attached to base stations, number of base stations, handover procedures, routing, security scanning etc. For consistency, a single UE is running on two base stations. Performing the necessary evaluations is enacted without changing their state, i.e. power-saving, due to the UE device being left in the premises of the Oslo Metropolitan University, and the impossibility to access the perimeter because of the Covid-19 pandemic crisis (Brzozowski, 2020).

Table 7. CPU utilization summary of the 5GC core and the Centralized Unit physical nodes during a single-stream transmission test.

CPU utilization (%) 5GC_total	CPU utilization (%) 5GC_user	CPU utilization (%) 5GC_system	CPU utilization (%) CU_total	CPU utilization (%) CU_user	CPU utilization (%) CU_system
61.9658	8.23386	53.7319	35.983	5.03469	30.9483

The same rule applies for the Centralized Unit node, which is the sender in this case. The Total CPU utilization during the single-stream test is 35.983% and the current OS user is spending 5.03469% of the total amount for performing the test. Nevertheless, the overall CPU utilization by the OS and the Centralized Unit processing is taking 30.9483% from the available CPU resources in a dual-CPU configuration. The physical underlying hardware node is Dell PowerEdge R620 server (Dell Inc., 2012) with dual Xeon E5-2600 CPUs and 768GB RAM memory. The operating system is running on top of a RAID-0 configuration for maximizing read/write performance and without mirroring, allowing for maximum performance from the node. The Centralized Unit operations are substantially intensive in terms of computational resources demand, and therefore we allow for dual CPU architecture, where the aforementioned test considers the total available computational capacity.

4.1.2. Multiple-stream hardware-level tests

For the staple of convenience and attaining a better insight in the link capacity and performance, we perform a parallel stream test with 10 different streams, each occupying a different socket. To achieve that, we pass an argument *-P 10* to the command, allowing for 10 streams to run in parallel and equally share the bandwidth of the link. The TCP buffer size is equivalent, as well as the runtime of 300 seconds with transmissions triggering every consecutive second.

```
iperf3 -t 300 -i 1 -w 32M -c 10.0.0.1 -P 10
```

The results are similarly consistent to the tests with a single stream (see Figure 61 and Figure 62) and are summarized in Table 9. Notably, on socket 19, a retransmission occurs due to the link capacity being overburdened and the sender (the Centralized Unit) attempts a new TCP transmission. This however does not affect the overall performance, nor adds to the error rate significantly. Increasing the TCP buffer and/or MTU size on the interfaces may aid in that aspect, but that may also incite additional CPU overhead on both sides (sender – CU and receiver – 5GC, where the server is running).

Multiple-stream tests allow for a clear overview of the architecture and whether it can scale appropriately. Comparing these results with the further tests on the virtualization plane and the VPN tunnels, will ascertain a performance detriment in terms of hardware impact from the softwareization and virtualization. In parallel, this directly translates to performance impacts on the instantiated network slices in the 5G network.

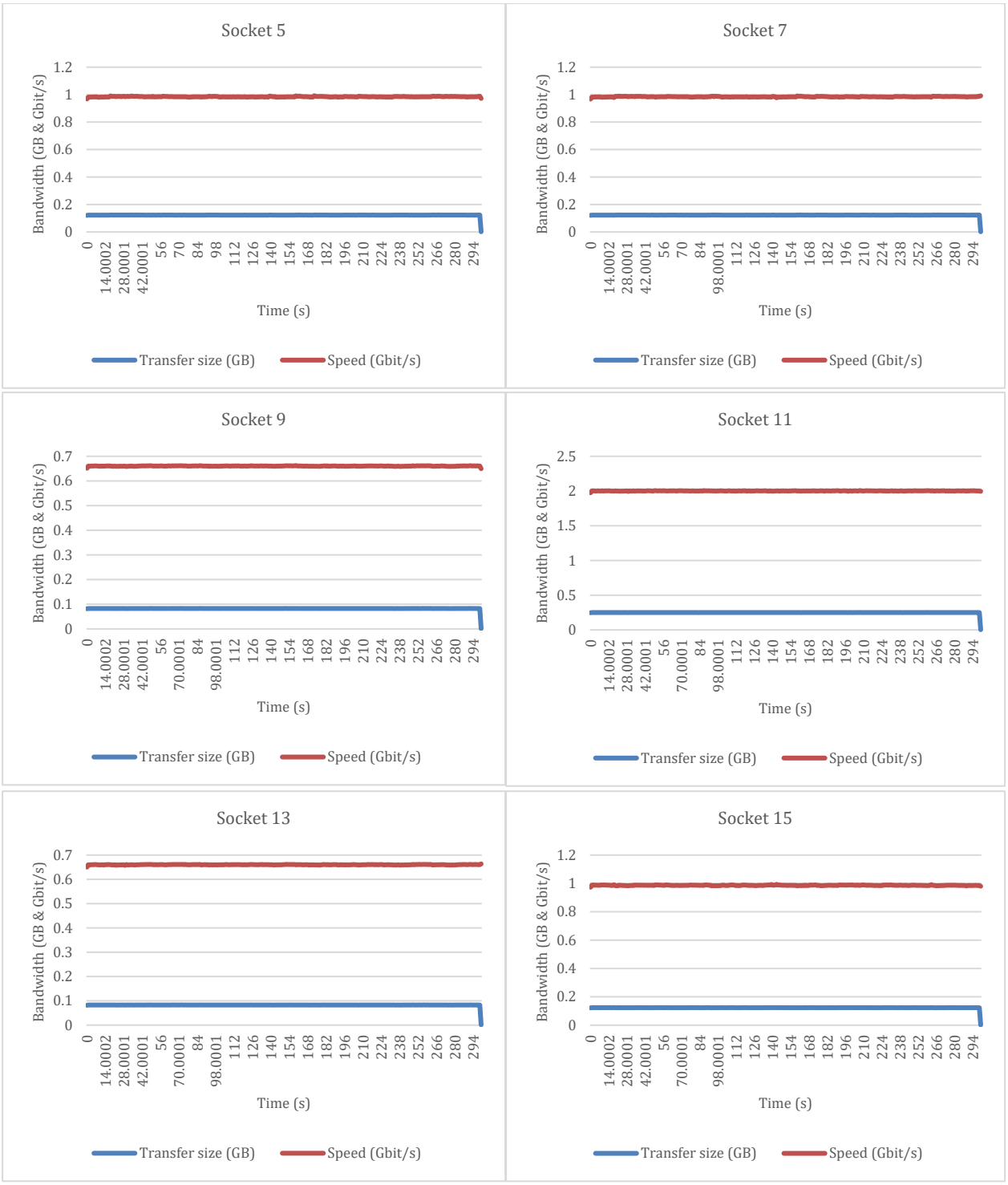

Figure 61. Parallel stream test (first 6 streams, sockets 7-15), sharing the total bandwidth of the physical link

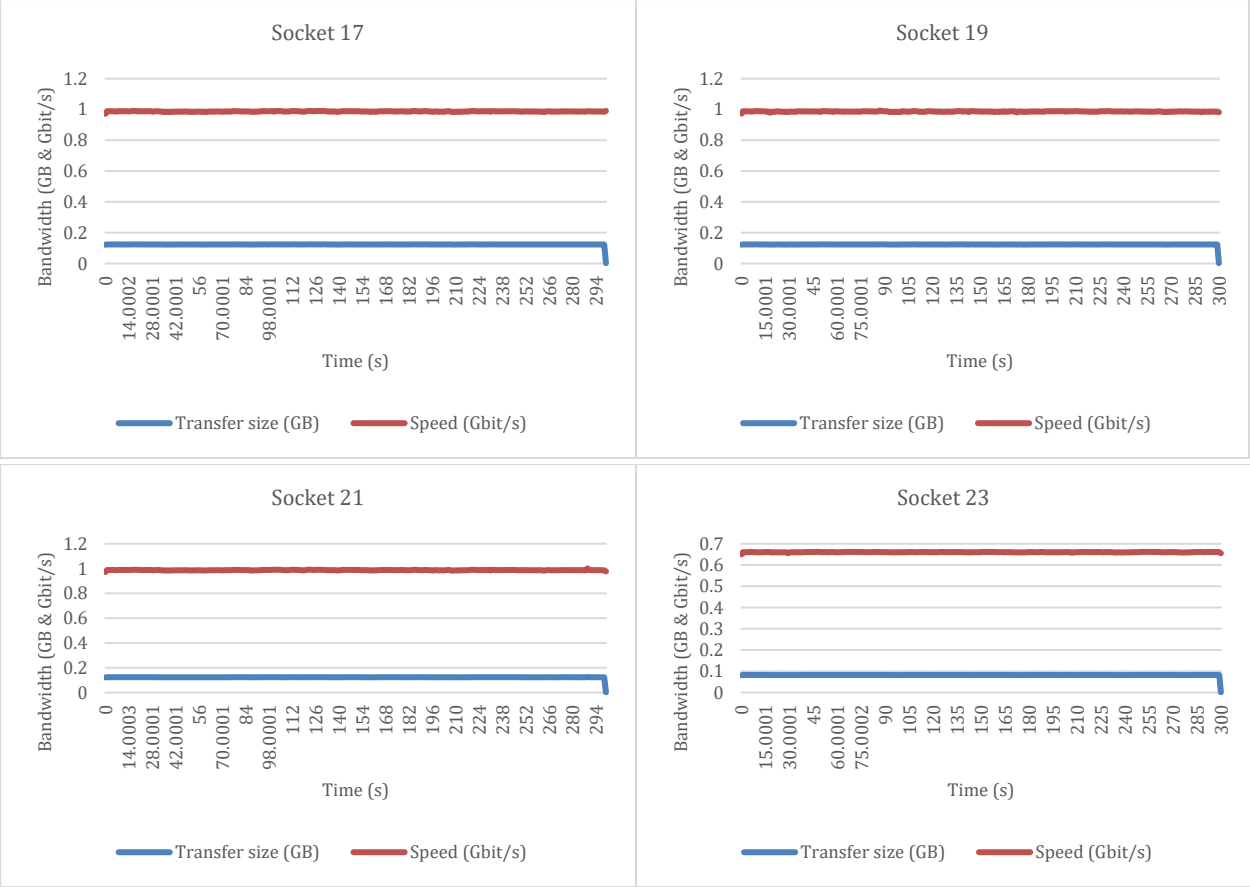

Figure 62. Parallel stream test (last 4 streams, sockets 17-23), sharing the total bandwidth of the physical link

The complete amount of data transmitted equals to 3711696551 Bytes, or 371.169 Gigabytes. The average maximum speed for all the streams in total is equal to 9.89736 Gbit/s, which does not differ from the single-stream test. This equilibrium however extends at a cost of computational resources, which the network attains by increasing the processing drastically for systems with less available resources.

Table 8. Summary of the parallel stream tests on the hardware network

Send Start (s)	Send Duration (s)	Sent Data (GB)	Send Speed (Gbps)	Retr. (total)	Rec. Start (s)	Rec. Duration (s)	Rec. Data (GB)	Rec. Speed (Gbps)
0	300.015	371.1696551	9.89736	**1**	0	300.015	371.1654621	9.89725

Table 9. Summary of the parallel stream test with 10 simultaneous streams, each sharing equal amount of bandwidth in the total link capacity.

End streams	Socket	Start (s)	Stream end time (s)	Stream duration time(s)	Total stream data (GB)	Total stream bandwidth (Gbit/s)	Retrans. (total)
sender	5	0	300.015	300.015	36.91105172	0.984245	0
receiver	5	0	300.015	300.015	36.91062221	0.984234	
sender	7	0	300.015	300.015	36.90241906	0.984015	0
receiver	7	0	300.015	300.015	36.90203253	0.984005	
sender	9	0	300.015	300.015	24.76341507	0.660324	0
receiver	9	0	300.015	300.015	24.76297662	0.660313	
sender	11	0	300.015	300.015	75.05832376	2.00145	0
receiver	11	0	300.015	300.015	75.0579211	2.00144	
sender	13	0	300.015	300.015	24.76253059	0.660301	0
receiver	13	0	300.015	300.015	24.76209214	0.660289	
sender	15	0	300.015	300.015	36.99828214	0.986571	0
receiver	15	0	300.015	300.015	36.9978777	0.98656	
sender	17	0	300.015	300.015	37.01184562	0.986933	0
receiver	17	0	300.015	300.015	37.01140717	0.986921	
sender	19	0	300.015	300.015	36.98974946	0.986344	1
receiver	19	0	300.015	300.015	36.98931101	0.986332	
sender	21	0	300.015	300.015	37.01003144	0.986884	0
receiver	21	0	300.015	300.015	37.00960912	0.986873	
sender	23	0	300.015	300.015	24.7620062	0.660287	0
receiver	23	0	300.015	300.015	24.76161249	0.660276	

As noted in Table 10, the total CPU utilization at the 5GC core network node is 95.9105%, from which the current user running the iPer3 server exploits 18.9873%. The system in parallel requires more resources to run, and thus increases the utilization of the CPU on 76.9232%. Similar trend is exhibited at the Centralized Unit node, where the overall CPU usage increments to 41.7224%, from which the user performing the test is utilizing 8.2758% in the current session. Additionally, the system CPU utilization increases as well, but not as with much offset as the Core Network, accounting for 33.4466% in total to support the operating system.

Table 10. CPU utilization summary of the 5GC core and the Centralized Unit physical nodes during 10 parallel streams transmission test.

CPU utilization (%) 5GC_total	CPU utilization (%) 5GC_user	CPU utilization (%) 5GC_system	CPU utilization (%) CU_total	CPU utilization (%) CU_user	CPU utilization (%) CU_system
95.9105	18.9873	76.9232	41.7224	8.2758	33.4466

4.2. Testing the virtualization-plane with SR-IOV drivers

In a similar manner like the hardware-level test, the virtualization layer assessment is performed with the SR-IOV interfaces engaged and both the Core Network and Centralized Unit analyzed. The tests are performed in two stages, using a single transmission queue as well as 10 parallel streams, consequently. The same iPerf3 commands are used for both cases, correspondingly to the previous tests. Furthermore, the resource utilization of the hardware is measured and compared between the scenario of a single-stream test and using parallel streams. For precise assessment, we obtain the results relevant for the virtualization plane only, which indicates that the numbers obtained are not considering the overall utilization of the whole underlying hardware.

4.2.1. Single-stream tests on the SR-IOV virtualization-plane

As seen in Figure 63, the bandwidth over 300 seconds is identical as the tests performed directly on the hardware. The sender reports 9.89704 Gbit/s speed through the SR-IOV drivers, compared to the hardware-level tests where the sender achieved 9.89328 Gbit/s. This rather small difference is in the margin of error and we can conclude that the SR-IOV does not add to any network overhead (see Table 11). This, however, comes at a hardware-level cost in terms of resources. For the SR-IOV drivers to attain these hardware-level network performances, there is a slight sacrifice in terms of computational resource capacity.

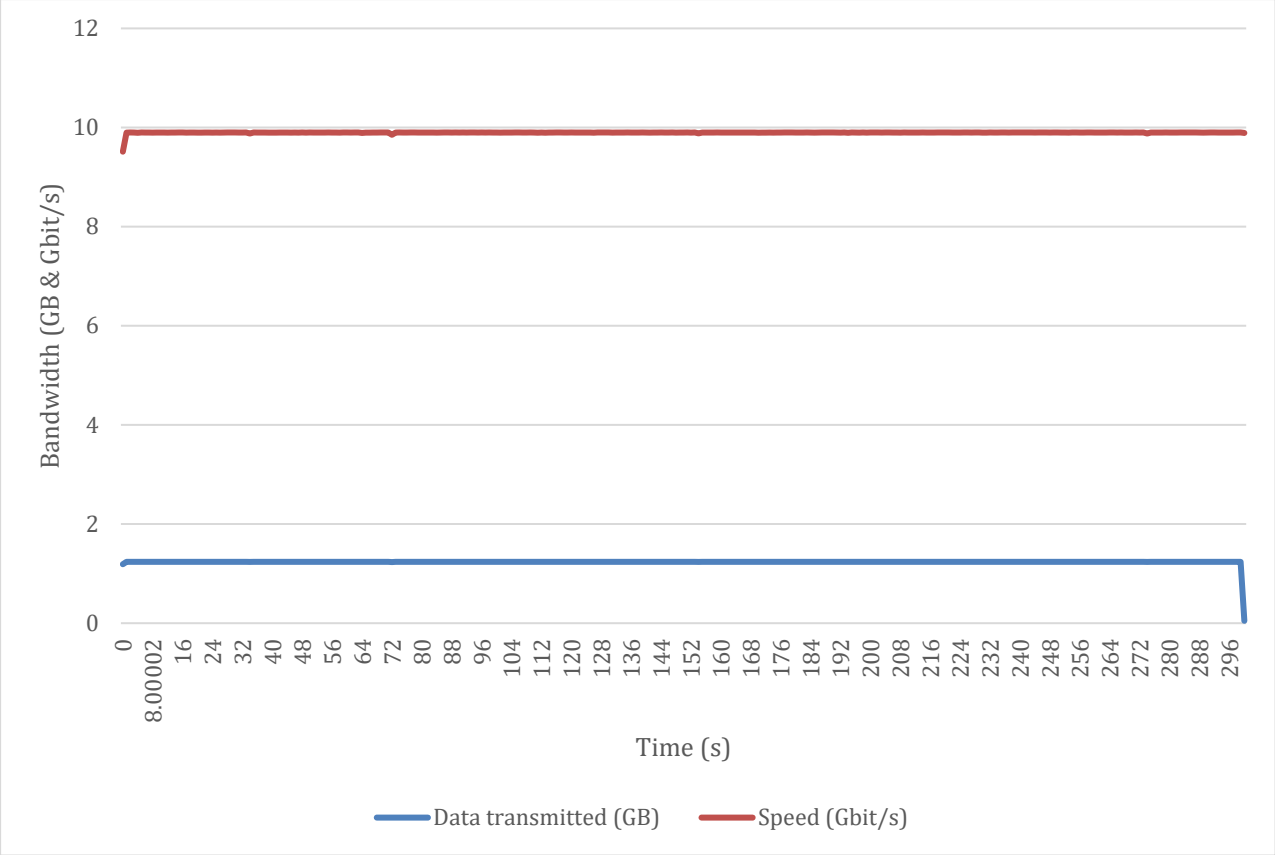

Figure 63. Single-stream transmission performance between the Core Network and the Centralized Unit over the virtualization plane using SR-IOV.

Table 11. Summary of the single-stream tests performed on the virtualization-plane through the SR-IOV drivers. The results account the SR-IOV performance only. Hardware-level tests are excluded in these numbers.

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (GB)	Average Stream Bandwidth (Gbit/s)	Retrans. (total)
sender	5	0	300.039	300.039	371.187164	9.89704	0
receiver	5	0	300.039	300.039	371.1868419	9.89703	

The results of the total CPU utilization on the system is measured and assessed during the single-stream tests. There is a slight increase in the CPU utilization when SR-IOV is implemented at the server-level, which is the Core Network. It should be noted that the Centralized Unit node is more computationally capable, and any offloading will result in improvements of its performance. The case with the total CPU overhead incurred by SR-IOV at

the server node is 1.605%, whereas the CU experiences relief in the system processes and reports drop from 30.9483% onto 17.9793%, which is a drop of 12.969%. This dynamicity in the Linux kernel is taken into account and measured on the total CPU utilization, after which the results are subtracted in order to obtain the level of impact the SR-IOV drivers have on the overall CPU performance.

Initially, the overall CPU utilization on the hardware level was 30.9483%, from which we subtract the current 21.4606% and determine that the reported difference is 14.5224%. To obtain the realistic view on the situation and how much the SR-IOV drivers exhibit an impact on the performance, it is necessary to eliminate the system resource utilization, which was 12.969%. Therefore 14.5224 – 12.969 is 1.5534%. With this, we can conclude a consistency between the results at the Centralized Unit and the Core Network server node, which indicates that the SR-IOV drivers have 1.55% to 1.6% additional overhead on the CPU resource utilization, within the margin of error.

Table 12. SR-IOV resource utilization, together with the available hardware-level resources. The numbers show the total used resources on the physical machine.

CPU utilization (%) 5GC-SRIOV_total	CPU utilization (%) 5GC-SRIOV_user	CPU utilization (%) 5GC-SRIOV_system	CPU utilization (%) CU-SRIOV_total	CPU utilization (%) CU-SRIOV_user	CPU utilization (%) CU-SRIOV_system
63.5708	8.62531	54.9455	21.4606	3.48127	17.9793

4.2.2. Parallel stream tests on the SR-IOV virtualization-plane

Following are the tests on the SR-IOV virtualization plane during 10 parallel streams. The same method is applied with TCP congestion window buffer of 32 Megabytes, 300 seconds (5 minutes) test, with transmission every one second. The results of the 10 parallel streams bound to adjacent sockets are depicted in Figure 64 and Figure 65.

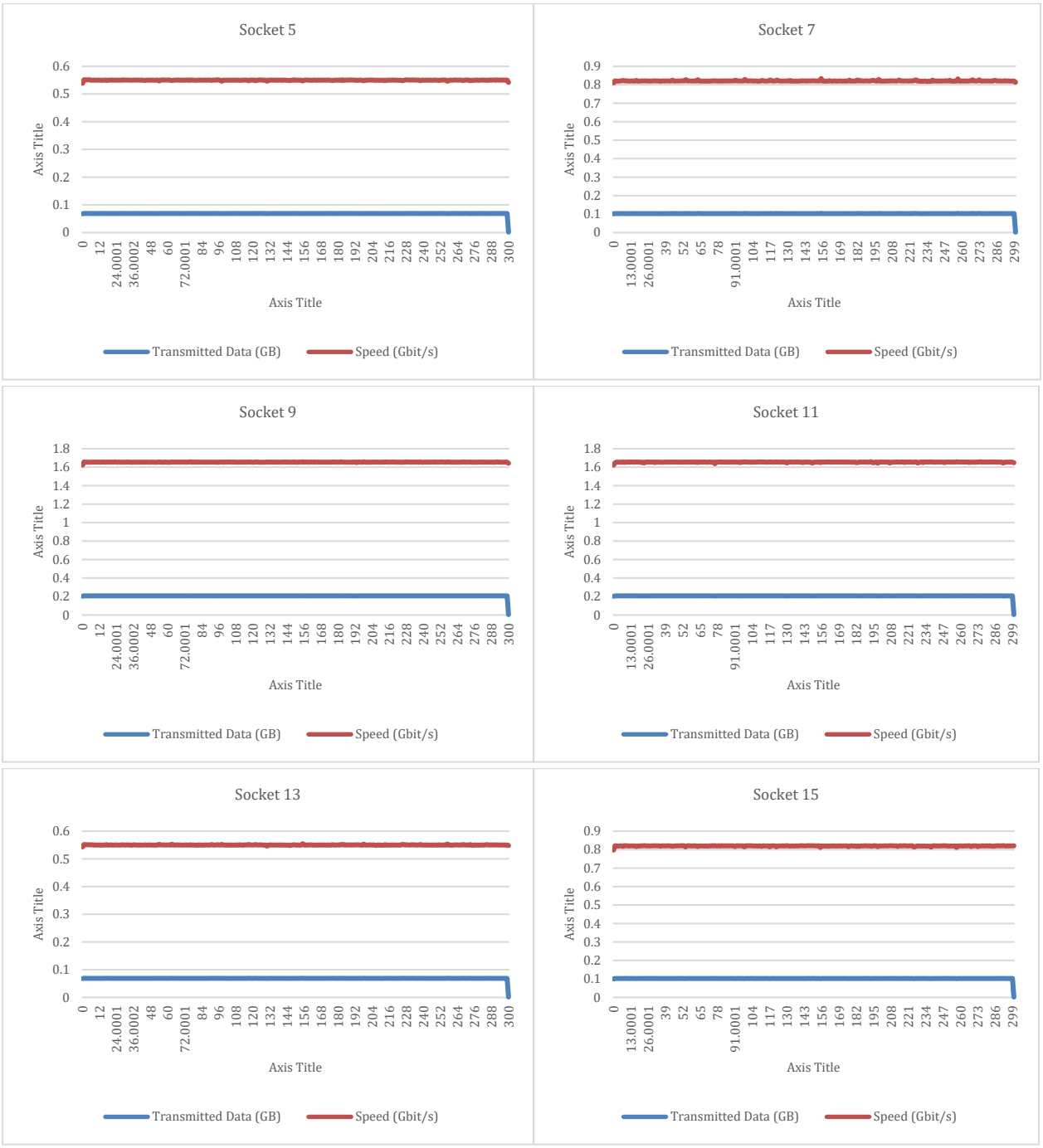

Figure 64. First 6 sockets of the SR-IOV virtualization-plane tests with 10 parallel streams and identical TCP buffer of 32 MB as in the previous tests. The results prune in time of 300 seconds, with transmissions occurring every 1 second.

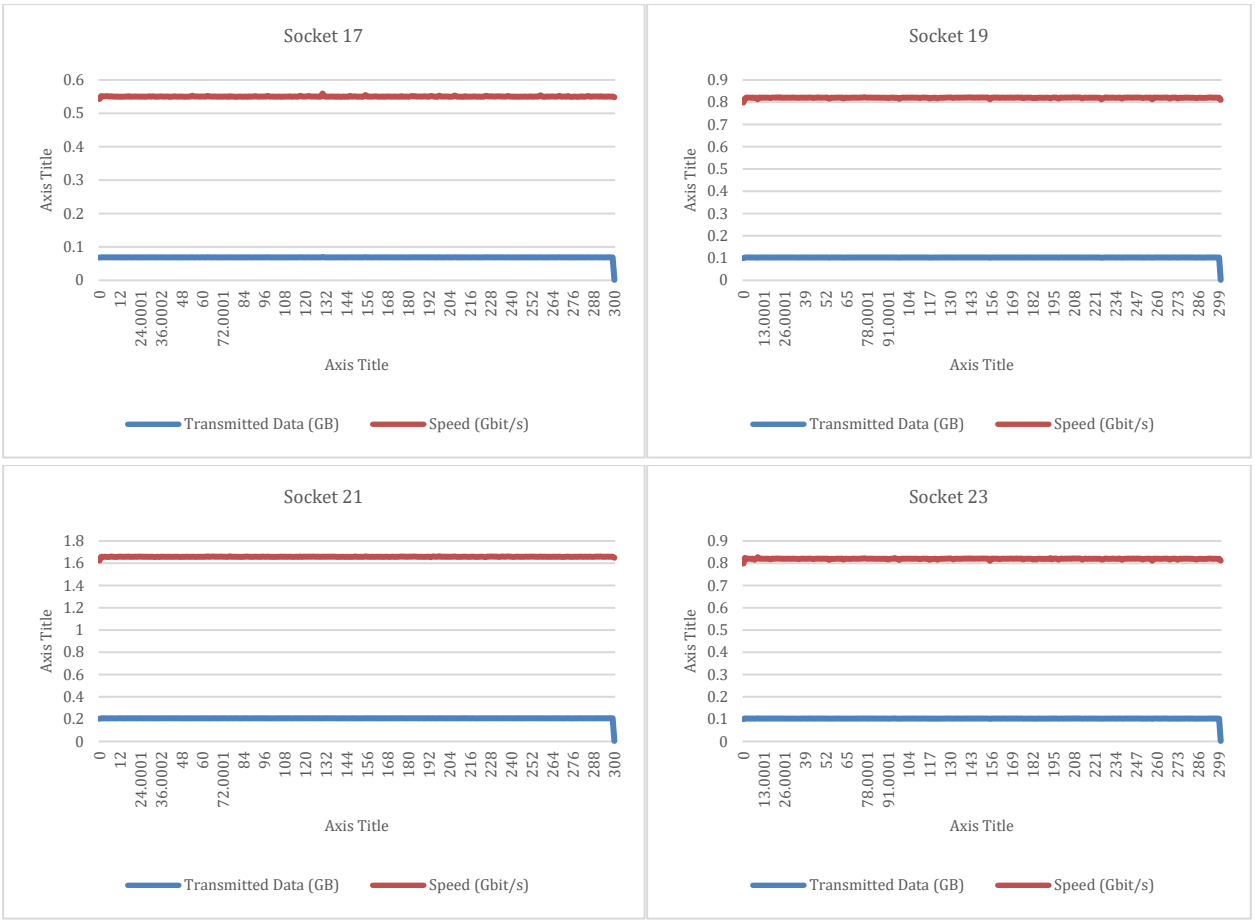

Figure 65. Last four sockets of the parallel transmission stream test on the SR-IOV virtualization-plane.

Table 13 summarizes the performance of each socket during the parallel stream test. Notably, retransmissions start to occur, which inclines towards the fact that scaling in the virtualization plane can represent a problem if there is only one virtual function (VF) engaged, as in the case of the tests in order to retain consistency. The redistribution of data between the sockets is not absolutely equal and therein the difference of speed in each socket. Nevertheless, this does not represent any performance issue for the overall transmission achieved.

Table 13. Summary of the SR-IOV virtualization-plane tests with 10 parallel streams. The representation is based on a single socket, showing the sender's and receiver's achieved bandwidth over 300s.

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (GB)	Average Stream Bandwidth (Gbit/s)	Retrans. (total)
sender	5	0	300.021	300.021	20.61880747	0.549796	2
receiver	5	0	300.021	300.021	20.61838691	0.549785	
sender	7	0	300.021	300.021	30.77852193	0.820703	0
receiver	7	0	300.021	300.021	30.7781175	0.820692	
sender	9	0	300.021	300.021	62.05316887	1.65464	0
receiver	9	0	300.021	300.021	62.05284674	1.65463	
sender	11	0	300.021	300.021	62.04547876	1.65443	8
receiver	11	0	300.021	300.021	62.04514768	1.65442	
sender	13	0	300.021	300.021	20.63077724	0.550116	1
receiver	13	0	300.021	300.021	20.63040142	0.550105	
sender	15	0	300.021	300.021	30.74309818	0.819759	2
receiver	15	0	300.021	300.021	30.74272236	0.819748	
sender	17	0	300.021	300.021	20.63407905	0.550204	0
receiver	17	0	300.021	300.021	20.63370324	0.550194	
sender	19	0	300.021	300.021	30.74551059	0.819823	3
receiver	19	0	300.021	300.021	30.74512583	0.819813	
sender	21	0	300.021	300.021	62.18255765	1.65809	0
receiver	21	0	300.021	300.021	62.18223552	1.65808	
sender	23	0	300.021	300.021	30.74611906	0.819839	6
receiver	23	0	300.021	300.021	30.74574324	0.819829	

Evidently, in Table 14, there are total of 22 TCP retransmissions for maintaining the maximal link capacity in the virtualization plane. In order for this to pertain, SR-IOV increases the demand for computational resources and therein the adjacent increment in CPU utilization in Table 15.

Table 14. Summary of the total transmission performance during parallel stream tests on the SR-IOV virtualization-plane, at the sender and receiver, correspondingly.

Sender Start (s)	Sender Duration (s)	Sent data (GB)	Send Speed (Gbps)	Retr. (total)	Rec. Start (s)	Rec. Duration (s)	Rec. data (GB)	Rec. Speed (Gbps)
0	300.021	371.1781188	9.89739	**22**	0	300.021	371.1744304	9.89729

Table 15. Total resource utilization obtained from the SR-IOV virtualization-plane tests, performed during parallel stream tests.

CPU utilization (%) 5GC-SRIOV_total	CPU utilization (%) 5GC-SRIOV_user	CPU utilization (%) 5GC-SRIOV_system	CPU utilization (%) CU-SRIOV_total	CPU utilization (%) CU-SRIOV_user	CPU utilization (%) CU-SRIOV_system
96.2447	20.1134	76.1313	27.36	4.43466	22.9254

In Table 15, there is a clear evidence of the impact of utilizing the virtualization plugin for Docker. Contrary to the overall CPU usage in the physical hardware tests with multiple streams, which was 95.9105%, at this point we have increase of 0.3342%. This, however, can have inconsistencies in terms of system processes variations, and thereby we compare the system CPU utilization during hardware-level tests and the SR-IOV virtualization-plane tests. During the hardware-level tests, there were 76.9232% of used resources, compared to the diminished usage during the SR-IOV tests, which is 76.1313%. The delta offset in this situation is 0.7919%, which we add to the 0.3342% to measure the impact the SR-IOV virtualization has on this multi-stream scenario at the Core-Network 5GC server node. Consequently, the overall impact of the virtualization is 1.12%, which is marginally less than the 1.5-1.6% during the single-stream tests. These numbers may be insignificant, and they clearly do not signify any influence on the performance of the physical and logical networks, nevertheless, the retransmissions indicate that if scaled further over the same virtual function (VF), the performance may start degrading further.

At the Centralized Unit node, the performance difference is also clear. During the hardware-level tests, the overall utilization of the CPU in the node was 41.7224%, whereas during the SR-IOV parallel stream tests, the total usage amounts for 27.36%. Nevertheless, there is variation in the system processes at the different times of testing, for which if we calculate the offset, we get 33.4466% during hardware-level tests and 22.9254% during SR-IOV tests, results in $\Delta_{SR-IOV(m)}$=10.5212%. This value is subtracted from the overall performance in order to reduce the additive value and isolate the effect SR-IOV virtualization has during the multiple stream tests, that is 41.7224% during the hardware-level tests and 27.36% during the SR-IOV tests. The difference of 14.3624% subtracts the system resources difference of 10.5212%, obtaining 3.8412% impact on the virtualization plane during the SR-IOV tests with multiple streams.

This indicates a trend in increasing overhead when scaling up the network traffic over the same Virtual Function. Without instantiating additional virtual functions, it is likely that the impact would be more than linear over a threshold of difference between the system's available resources and the number of virtualized functions for the SR-IOV drivers, accounting for the traffic generated also in the 5G network on top of the diverse network slices. Furthermore, as the number of retransmissions increase, the level of computational resources required for retaining maximal bandwidth will increase, adding on the stress of the physical nodes. This further increases with adding an encrypted VPN tunnel, which will be shown in the subsequent tests.

4.3. Direct VPN+ connection testing

Comparable to the tests of the hardware and virtual-network layers, the VPN tunnels are examined in terms of performance impact they exhibit on the hardware network as well as through the virtualization plane. It is expected that the encryption performed on top of the same hardware shall incur tremendous performance cuts, and therefore, we calculate only the total resource utilization of the VPN infrastructure, eliminating the utilization on a system level and SR-IOV virtualization level. This is because the machines tend to reach 99.9% resource utilization and thus it is important to measure the level of overhead the encryption and tunneling add on the underlying transport network.

Furthermore, we run the 5G infrastructure on top of the Docker containers (see Appendices A), B) and C) for configuration of the tunnel and the overall network representation).

4.3.1. Single-stream tests of the VPN hardware-level tunneling

The initial tests are performed using a single stream for 300 seconds time and transmissions occurring every one second. The same TCP buffer size is used, that is 32 MB, in order to pertain to the previous and subsequent tests. As indicated in Figure 66, the performance degradation on hardware-level is tremendous. Furthermore, the traffic varies greatly, with drops in speed and packet size. It should be noted also that optimizations of any kind are not being performed in order to maximize the throughput and diminish the AES-256 encryption influence on the hardware. Adjusting the physical interfaces to support bigger packet size and encapsulation, maximizing the MTU and block sizes can improve the performance, but that is not the goal of

this thesis. We show the feasibility of successful implementation of advanced isolation of network slices in 5G using VPN on the transport network, and the performance degradation is shown in Figure 66, summarized in Table 16.

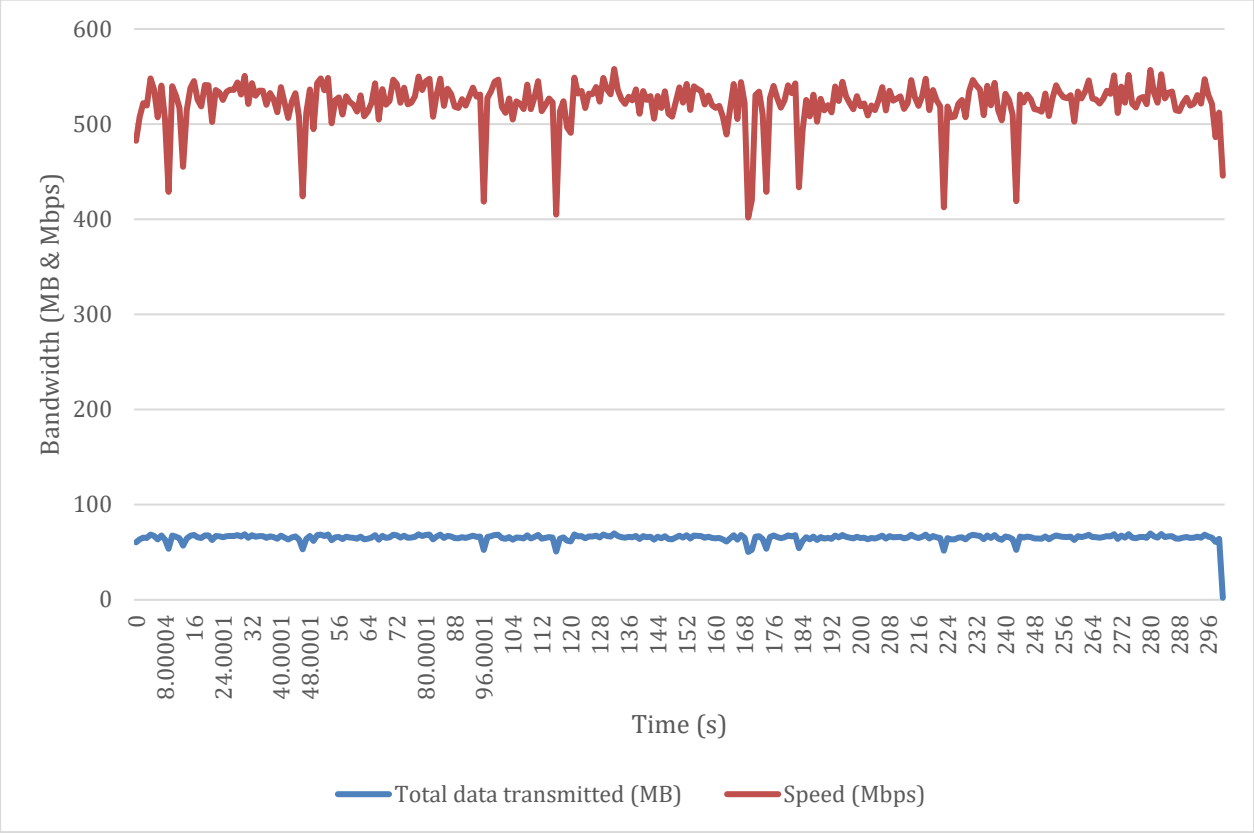

Figure 66. Encrypted VPN tunnel tests on the underlying infrastructure, connected between the two hardware endpoints directly using a single-stream test during 5 minutes time.

The performance degradation is immense, and it mostly accounts for the encryption of each packet, performed at the same hardware. This saturates the available computational resources otherwise reserved for the system processes and consequently, running the 5G core network and the centralized unit. Appendix C) denotes the Centralized unit communicating with the Core Network and the established GTP tunnels are represented.

Table 16. Summary of the single-stream tests on the VPN deployment in the direct link between the two instances (Core Network and Centralized Unit).

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (MB)	Average Stream Bandwidth (Mbit/s)	Retrans. (total)
sender	5	0	300.037	300.037	19591.88627	522.386	163311
receiver	5	0	300.037	300.037	19591.43266	522.374	

Compared to the hardware-level tests, the VPN tunneling eliminates good portion of the overall available bandwidth of the link. Namely, the drop amounts for 1,893.86%, from 9893.28Mbit/s to 522.386Mbit/s at the sender's side. In other words, the 10Gbit link drops to a 500Mbps link, which is orders of magnitude lower than the underlying network with or without the SR-IOV virtualization. The resource utilization is represented in Table 17. From all the available resources, the VPN processing utilizes 37.1798% at the server side and only 3.53544% at the client side, indicating a substantial difference between the necessity for computational resources for performing encryption of every packet passing in the tunnel. Due to insufficient power, the physical node at the Core Network is incapable of attaining the required maximal bandwidth, which results in 163311 retransmissions (Table 16).

Table 17. VPN tunneling resource utilization calculated distinctly from the overall system utilization in the underlying hardware (on the Core Network and Centralized Unit nodes, correspondingly).

CPU utilization (%) 5GC-VPN_total	CPU utilization (%) 5GC-VPN_user	CPU utilization (%) 5GC-VPN_system	CPU utilization (%) CU-VPN_total	CPU utilization (%) CU-VPN_user	CPU utilization (%) CU-VPN_system
37.1798	10.0446	27.1352	3.54544	0.36947	3.17597

4.3.2. Multiple-stream tests of the VPN hardware-level tunneling

The effect of performance degradation becomes more evident as the traffic is scaling and parallel transmissions are introduced. Without a proper FPGA-based hardware offloading for encryption computations, the stress on the CPU is additive and the machine needs to forfeit from the system processes to reduce their priorities, as well as the network controller, in order to manage the encryption and successful transmission. In this case, however, we witness not only increased retransmissions, but also total packet drops (transmission does not occur). The variations in the traffic at each socket is represented in Figure 67 and Figure 68.

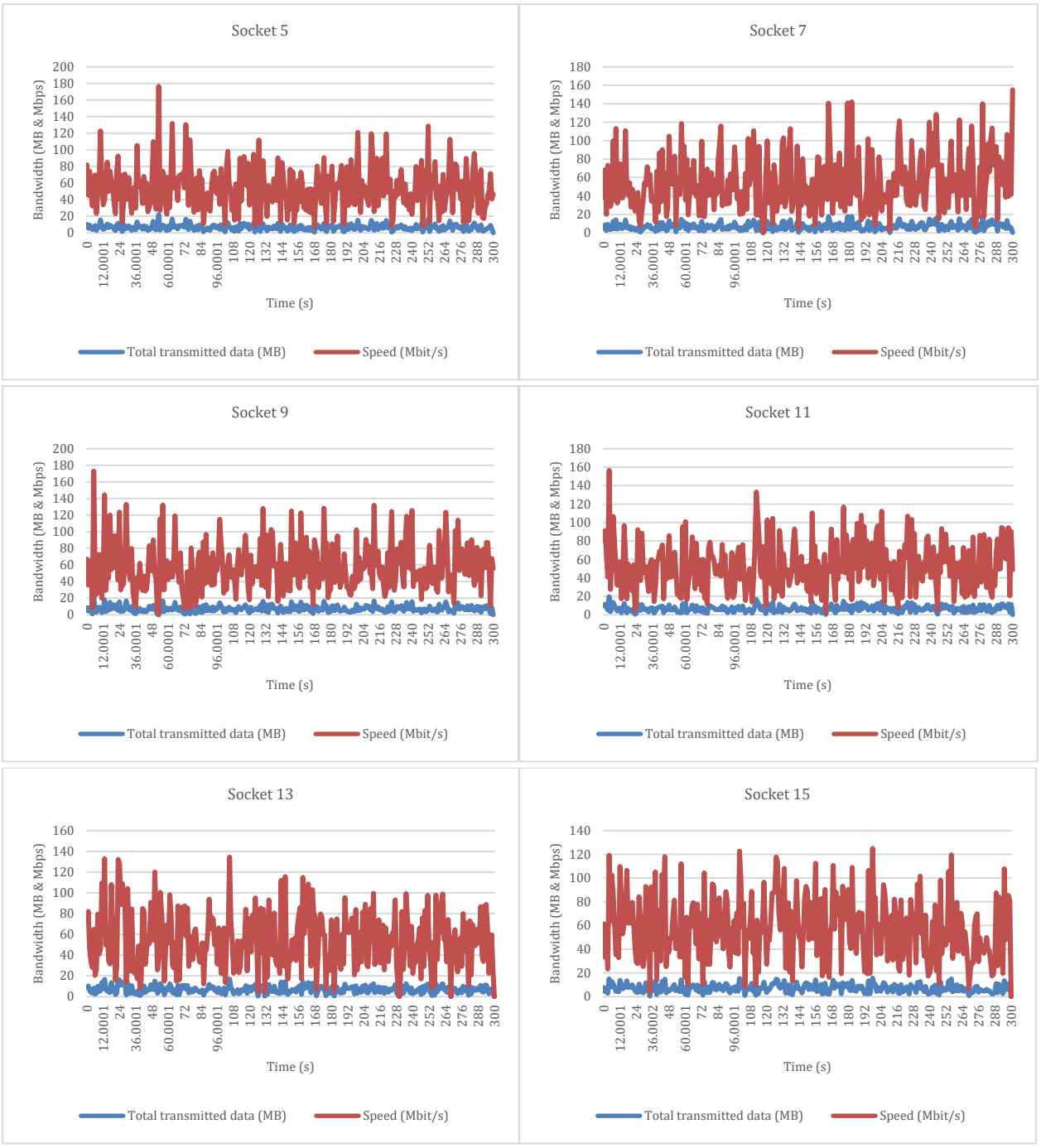

Figure 67. Multiple-stream transmission test of the underlying VPN connection on the hardware layer. The first 6 socket streams are depicted.

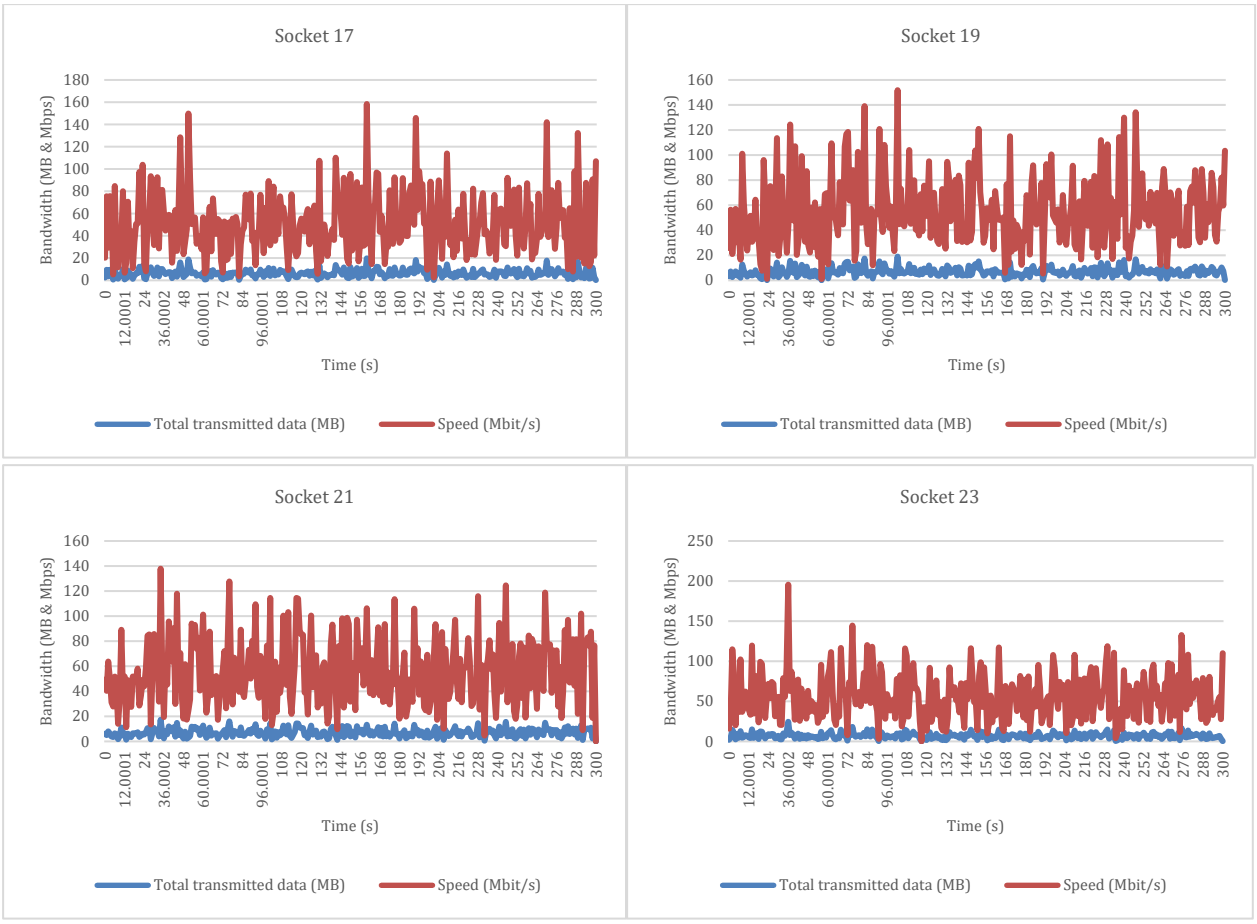

Figure 68. Last four socket streams from the multi-stream VPN hardware-level tests.

Table 18 summarizes the effects of the encryption on each socket. The number of retransmissions per socket is equally redistributed, showing a trend of overall exertion of the underlying hardware to sustain the connection during encryption. The packet loss is also much higher than when scaling the SR-IOV virtualization-plane without encryption.

Table 18. Summary of the multiple-stream transmission, indicating each socket performance and total rentrasmission attempts during VPN tunneling.

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (MB)	Average Stream Bandwidth (Mbit/s)	Retransm. (total)
sender	5	0	300.021	300.021	2017.309424	53.7911	34535
receiver	5	0	300.021	300.021	2016.845054	53.7787	
sender	7	0	300.021	300.021	2052.262352	54.7231	34575
receiver	7	0	300.021	300.021	2051.838362	54.7118	
sender	9	0	300.021	300.021	2156.66888	57.5071	36480
receiver	9	0	300.021	300.021	2156.257004	57.4961	
sender	11	0	300.021	300.021	1985.920704	52.9541	33107
receiver	11	0	300.021	300.021	1985.475178	52.9423	
sender	13	0	300.021	300.021	2057.81864	54.8713	33764
receiver	13	0	300.021	300.021	2057.366384	54.8592	
sender	15	0	300.021	300.021	2151.696756	57.3745	35009
receiver	15	0	300.021	300.021	2151.247192	57.3625	
sender	17	0	300.021	300.021	1895.523344	50.5437	32992
receiver	17	0	300.021	300.021	1895.085894	50.532	
sender	19	0	300.021	300.021	2054.846672	54.792	35583
receiver	19	0	300.021	300.021	2054.394416	54.78	
sender	21	0	300.021	300.021	2098.973936	55.9687	35480
receiver	21	0	300.021	300.021	2098.52841	55.9568	
sender	23	0	300.021	300.021	2092.448528	55.7947	35614
receiver	23	0	300.021	300.021	2092.032614	55.7836	

According to Table 19, the difference between the unencrypted communication and the encrypted VPN tunneling is drastic. Without proper optimization and additional computational entities, the deployment of a software-defined virtual tunneling transport network will employ significant traffic dilapidation.

Table 19. Overall performance summary of the VPN deployment in the 5G network at the Sender side (CU) and Receiver side (5GC).

Send Start (s)	Send Duration (s)	Sent Data (MB)	Send Speed (Mbps)	Retr. (total)	Rec. Start (s)	Rec. Duration (s)	Received Data (MB)	Rec. Speed (Mbps)
0	300.021	20563.46924	548.32	347139	0	300.021	20559.07051	548.203

The performance difference between the single-stream and multiple stream tests is insignificant, varying between 522.386Mbit/s during single-stream tests and 548.32Mbit/s during multiple-stream tests. The retransmission number is highly increased, from 163311 during the single-stream tests up to 347139 retransmissions during the multiple stream tests. These differences are significant and the comparable difference in speed between the both tests is within margin of error and variability. The same applies for the resource utilization, where the increase from 37.1798% during single-stream tests arises to ca% during multiple-stream tests. Incrementally, 7.0825% does not seem like a large sum. Nevertheless, this is 7% the VPN encryption is only cutting from the overall available CPU resources. This means that from the 55.7377% available computation power, the underlying hardware now has 48.6552% available resources for networking. The same will translate into higher impact on the instantiated network slices and the unpredictability of the traffic.

Table 20. Resource utilization of the VPN infrastructure during the multiple stream testing. The numbers show the exclusive CPU usage of the VPN deployment, without the total system utilization of other services.

CPU utilization (%) 5GC-VPN-TN_only	CPU utilization (%) 5GC-VPN-TN_user	CPU utilization (%) 5GC-VPN-TN_system	CPU utilization (%) CU-VPN-TN_only	CPU utilization (%) CU-VPN-TN_user	CPU utilization (%) CU-VPN-TN_system
44.2623	8.84595	35.4163	4.68069	0.944212	3.73648

The conclusive evidence suggests that instantiating a Software-Defined VPN transport network between the 5G Core Network and the Centralized Unit backhaul can have detrimental impact on the performance. Although the security is substantially strengthened, the communication may result in interruptions and inadequacy of utilization of real-time applications that require consistency in the bandwidth.

Although the aim of this work is not to improve the performance of VPNs and decide on traffic-steering mechanisms for attaining near-hardware level executions in the transport network slicing using VPN hard isolation, we measure the difference between the virtualized VPN+ SR-IOV 5GC core and Centralized Unit, with the previous tests of the tunneling on hardware-level. This shall provide better understanding whether the tunneling can be achieved and will have any impact related to the virtualized SR-IOV Docker deployment of the Core Network and the Centralized Unit in 5G.

4.4. Tests on the VPN+ transport network between the 5GC and CU within SR-IOV

To obtain a clear understanding whether the VPN has a substantial impact on the transport network backhauling between the 5GC and the CU, the results from the difference between the SR-IOV virtualization and the hardware-level performance are compared with the difference between the results from the following VPN SR-IOV Virtual Network Function results and the VPN hardware-level results. With this, we can ascertain a level of effect in the transport network without accounting for the encryption repercussions and the traffic degradation, which can be improved with additional methods such as employment of a FPGA-based fabrics for offloading the VPN-related functions.

Initially, we perform a single-stream test on the VPN isolation at the SR-IOV virtual function layer, and consequently a multiple-stream tests equal to the previous tests on the hardware-level and logical-level environments.

4.4.1. Single-stream tests of the VPN isolation at a SR-IOV VNF

The results from the multiple-stream tests are similar to the single stream tests in the VPN virtualized environment. This trend is equal to the difference between the performance at a hardware level and the virtualized environment using SR-IOV. According to Figure 69, the traffic experiences fluctuations as was in the case of direct VPN deployment on top of the bare-metal hardware.

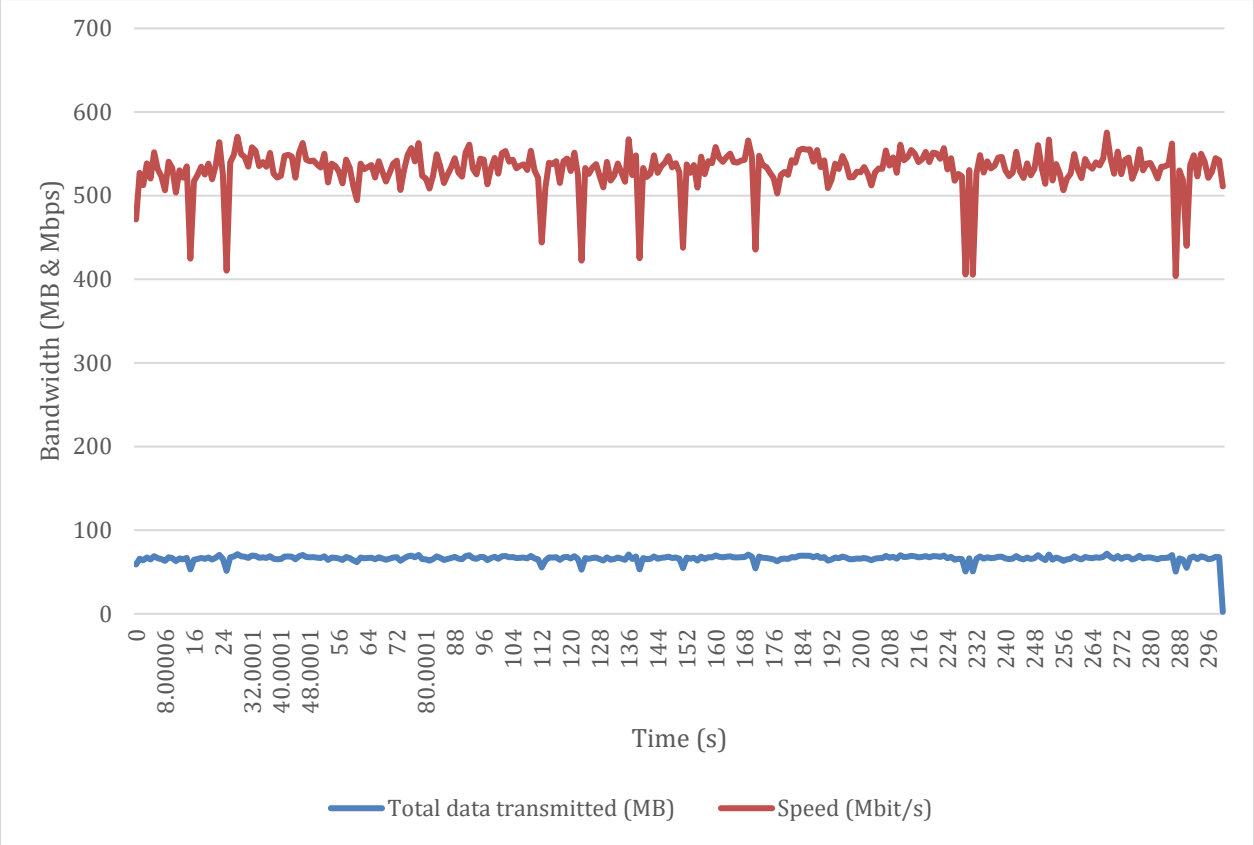

Figure 69. Software-Defined VPN single-stream tests over the SR-IOV virtualization plane between 5GC and CU.

The difference in performance is also insignificant, within a margin of error, indicating similar trend and relationship between the system processes performance forfeiting and the level of CPU utilization by the VPN exclusively.

Table 21. Summary of the overall transmission performance through VPN tunneling in the VNF of SR-IOV Docker instance

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (MB)	Average Stream Bandwidth (Mbit/s)	Retransm. (total)
sender	5	0	300.038	300.038	19923.19788	531.218	170510
receiver	5	0	300.038	300.038	19922.78466	531.207	

The number of retransmissions is higher in the virtualized SR-IOV VPN deployment, rising from 163311 to 170510. The network overall speed at the sender side is similar, with 522.386 Mbps in the hardware-level VPN tests, compared to the 531.218 Mbps in the transport network

tunnel between the SR-IOV endopints. The higher network performance also comes at a cost of a slightly increased hardware utilization. Namely, the difference between the system resource usage is subtracted with the overall VPN utilization, to observe the influence of the SR-IOV virtualized environment. During the hardware-level VPN instance, the overall receiver's system CPU utilization was 27.1352%, whereas in the time of measuring the VPN instance through the SR-IOV endpoints is 26.9148%. This offsets a 0.2207% relief, which when subtracted with the hardware-to-SR-IOV difference in the overall CPU utilization of 0.187% (37.1798% - 36.9928%), gives an overall impact of 0.0337% overhead, which is insignificant in terms of global performance.

Therefore, the difference between the received traffic speed of 531.207 Mbit/s, contrary to the hardware-level tunneling performance of 522.374 Mbps is minor and can encompass the margin of error together with the offset in systems performance relief from the background processes and the lesser VPN service resource utilization of 0.0337%.

Table 22. Overall CPU utilization during VPN tunneling within a SR-IOV virtualization Docker instance of the Core Network and the Centralized Unit, respectively.

CPU utilization (%) 5GC-VPN-TN_only	CPU utilization (%) 5GC-VPN-TN_user	CPU utilization (%) 5GC-VPN-TN_system	CPU utilization (%) CU-VPN-TN_only	CPU utilization (%) CU-VPN-TN_user	CPU utilization (%) CU-VPN-TN_system
36.9928	10.078	26.9148	1.89289	0.22533	1.66756

4.4.2. Multiple-stream tests of the VPN isolation at a SR-IOV VNF

Last but not least, the tests are scaled with the 10-parallels stream examination to determine the effects on the virtualized tunnel between the SR-IOV endpoints and the level of packet drops and traffic degradation in the 5G backhauling between the Docker instances of the 5GC and the CU. The results per each socket are depicted in Figure 70 and Figure 71. Notably, the packet drop level is increased slightly and the performance of transmission is similar to the hardware-level VPN tunneling test results.

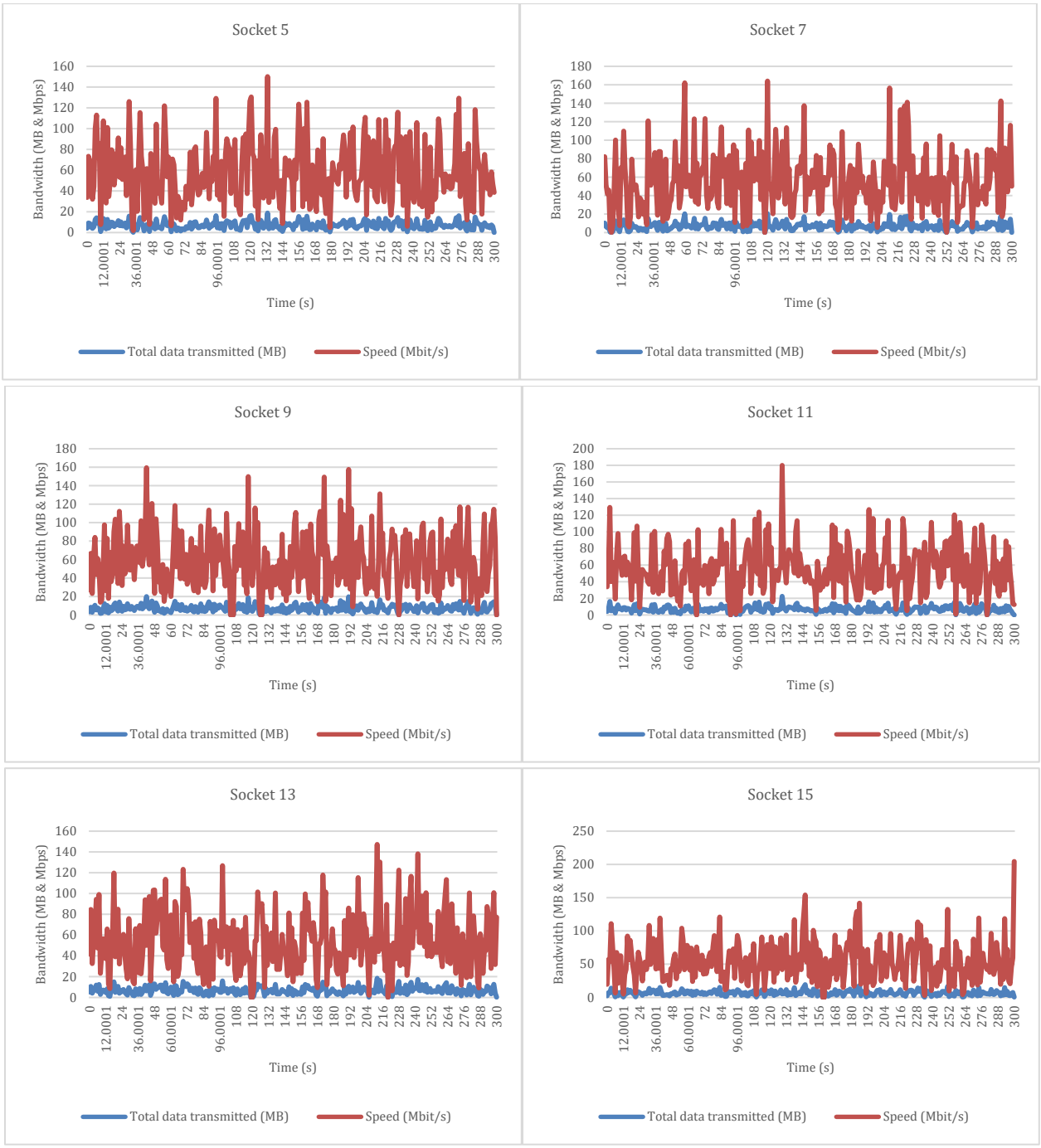

Figure 70. Multiple-stream test on the VPN tunnel between the two SR-IOV endpoints, the Core Network and Centralized Unit running in Docker (the first 6 sockets).

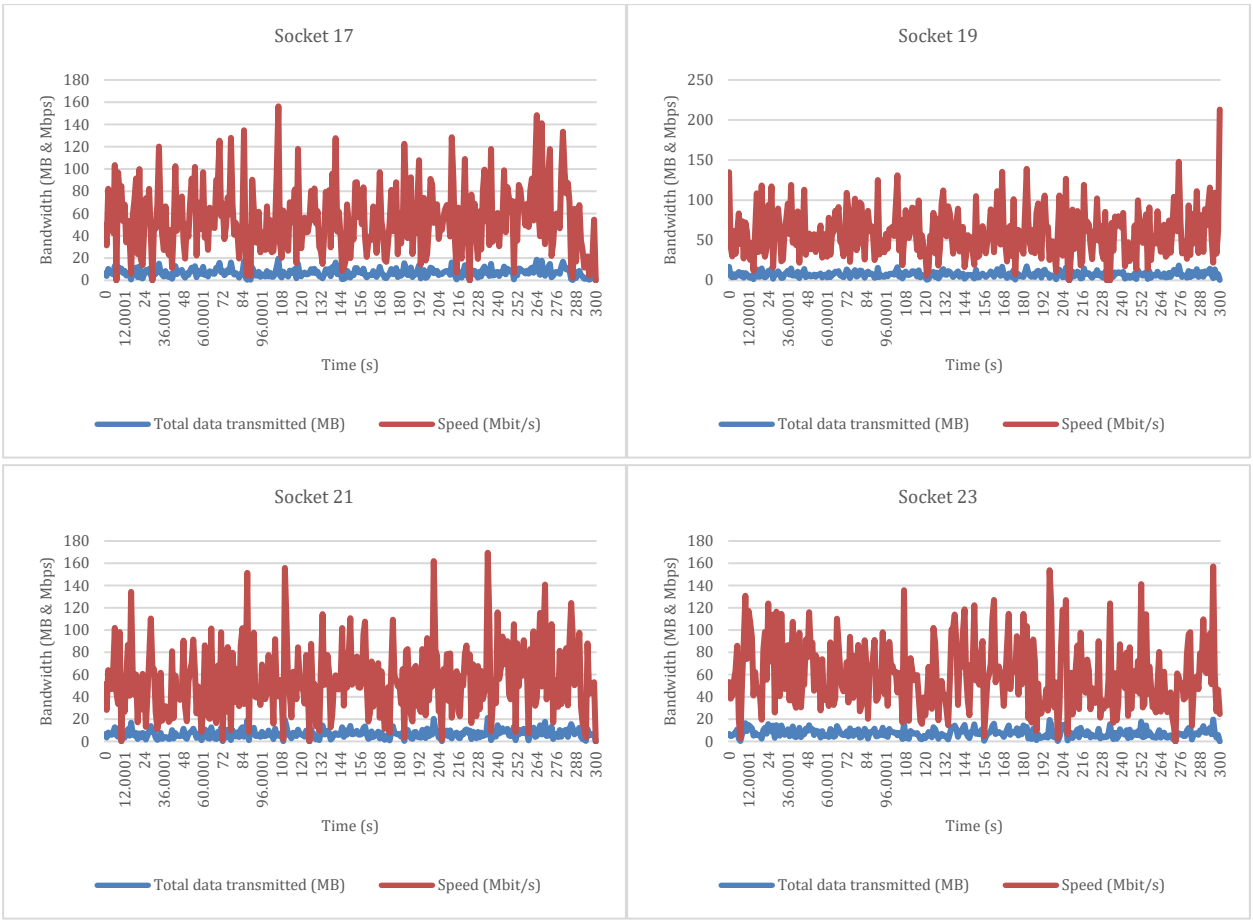

Figure 71. Multiple-stream test on the VPN tunnel between the two SR-IOV endpoints, the Core Network and the Centralized Unit running in Docker (the last 4 sockets).

The Table 23 sums the performance of each socket, which does not differ from the hardware-level VPN implementation and has elevated level of retransmissions. That implies on scaling problems and the potential severe traffic degradation in the VPN tunnel due to lack of computational resources. The fact that the VPN service only utilizes almost 45% of the whole server indicates that the virtualization endpoints are not representing any issue as much as the encryption encompasses in the physical machine.

Table 23. Summary of the test results during multiple-streams per each socket in a VPN tunnel between the two SR-IOV endpoints

End-Streams	Socket	Start (s)	Stream end time (s)	Stream duration time (s)	Total stream data (MB)	Average Stream Bandwidth (Mbit/s)	Retransm. (total)
sender	5	0	300.022	300.022	2160.416144	57.607	36924
receiver	5	0	300.022	300.022	2159.997538	57.5958	
sender	7	0	300.022	300.022	2067.574448	55.1314	33663
receiver	7	0	300.022	300.022	2067.130268	55.1195	
sender	9	0	300.022	300.022	2167.135376	57.7861	34903
receiver	9	0	300.022	300.022	2166.666968	57.7736	
sender	11	0	300.022	300.022	2094.515984	55.8497	33647
receiver	11	0	300.022	300.022	2094.069112	55.8378	
sender	13	0	300.022	300.022	2021.63816	53.9065	35397
receiver	13	0	300.022	300.022	2021.188596	53.8945	
sender	15	0	300.022	300.022	1995.988784	53.2225	33326
receiver	15	0	300.022	300.022	1995.547296	53.2108	
sender	17	0	300.022	300.022	2064.473264	55.0487	34830
receiver	17	0	300.022	300.022	2064.047928	55.0373	
sender	19	0	300.022	300.022	2189.683568	58.3874	35571
receiver	19	0	300.022	300.022	2189.216506	58.3749	
sender	21	0	300.022	300.022	2053.166864	54.7472	34122
receiver	21	0	300.022	300.022	2052.719992	54.7353	
sender	23	0	300.022	300.022	2235.232208	59.6019	36170
receiver	23	0	300.022	300.022	2234.796104	59.5903	

The relative speed is represented as slightly higher in Table 24 compared to the hardware-level VPN tunneling, and this accounts to the slight offloading at the sender's side in the system services. This can be regarded within the margin of error and variability of the system, which does not represent any significant difference in the real traffic performance.

Table 24. Summary of the total multi-stream performance tests performed on the VPN tunneling between the SR-IOV endpoints of the 5GC and CU in the transport network

Send Start (s)	Send Duration (s)	Sent Data (MB)	Send Speed (Mbps)	Retr. (total)	Rec. Start (s)	Rec. Duration (s)	Received Data (MB)	Rec. Speed (Mbps)
0	300.022	21049.8248	561.288	348553	0	300.022	21045.38031	561.17

The relative difference between the system CPU utilization at the hardware level tunneling was 35.4163%, compared to the 35.216% in the tunneling between the SR-IOV endpoints, which equals 0.2003%. Accounting for the overall difference in CPU utilization, which is 0.2452% (44.2623% - 44.0171%). If we differentiate these values, we get a 0.0449% of improvement over the SR-IOV virtualization layer and is insignificant. If we observe the sender's performances, it is evident that the Centralized Unit node experiences a slight respite in terms of system service CPU utilization, and that translates directly to the overall CPU utilization of the tunneling. Because the sender does not perform the encryption, but rather uses the VPN tunnel, the sending rate is elevated with increased retransmissions in parallel. That results in higher CPU utilization at the CU node and can account for the marginal increase in speed.

Nevertheless, this minor speed increment is insignificant in terms of the overall underlying link capacity, that is 10Gbps, therefore attaining the notion that the SR-IOV virtualization does not impact the performance with or without VPN tunneling. Furthermore, the Docker instances utilizing the SR-IOV plugin perform flawlessly and correspond to the hardware-level performances.

Table 25. Overall CPU utilization factor during VPN tunneling and AES-256 encryption between the two SR-IOV endpoints simultaneously. The numbers represent the exclusive VPN process utilization of the CPU resources.

CPU utilization (%) 5GC-VPN-TN_only	CPU utilization (%) 5GC-VPN-TN_user	CPU utilization (%) 5GC-VPN-TN_system	CPU utilization (%) CU-VPN-TN_only	CPU utilization (%) CU-VPN-TN_user	CPU utilization (%) CU-VPN-TN_system
44.0171	8.80109	35.216	3.11648	0.811096	2.30538

5. Discussion

Within this thesis, we demonstrated the successful implementation of a VPN+ transport network between the Centralized Unit of a 5G C-RAN and the Core Network in the TN. Consequently, the experiments shed light on the feasibility of achieving near-identical performance in the virtualization plane using the SR-IOV drivers for container networking, rendering the communication between containers in a same subnet and different physical hosts, similar to the communication via direct link between the physical machines. The VPN transport network however, validates an immense impact on the inter-slice networking performance, not due to the virtualization, but rather the computational complexity of the cryptographic operations that overload the physical underlying hardware and deplete the otherwise available resources used for backhaul networking between the 5G core and the Centralized Unit.

It should be noted that the SR-IOV plugin can be run with an argument *privileged=1* into the command, which will allow SR-IOV to enable access to trusted networks. This may be useful when modifications on Layer-2 are necessary in the underlay, and the containers can be redistributed through the local network using their physical MAC addresses while being able to do modifications on Layer-2. For better security, the containers should be run in an isolated environment, where SR-IOV ensures hardware offloading to disable containers to perform ARP spoofing attacks or sniffing in general. If services running in multiple containers are trusted, then running these in microservices model (using Kubernetes) may benefit from being executed in a non-isolated ecosystem, such as a single network slice or single VLAN range. Nevertheless, for inter-slice networking, it is a basic requirement for the slices to be isolated as much as possible.

The combination of hardware offloading, isolation using distinct PFs (Physical Functions) and VFs (Virtual Functions) as well as policy enforcement, provides substantial security level that most enterprises deploying 5G will consider. Nevertheless, in some instances such as critical infrastructure, where the sacrifice of network performance is not a problem, VPNs may prove a viable possibility to harden the isolation between network slices. For IoT slices that do not require high bandwidth and low latency, the enhanced VPN shall provide the best isolation while maintaining decent level of QoS. To further address the performance impact induced at the transport network, the enhanced VPN can be supplemented by traffic steering techniques, which will ensure that no substantial loss is experienced between the endpoints in the slices.

To improve the performance of a VPN on the transport network, it is necessary to move the cryptographic processing outside the Core Network servers. A solution advocating the implementation of FPGA-based processing for offloading VPN cryptographic computations, combined with specific traffic-steering techniques can return the connection performance as through a generic virtualization plane, which was shown that does not differ significantly from the direct connection. However, in many instances this cannot be possible for the Telco operators and they will decide the network slice isolation to be conducted based on the combination of policy enforcement and advanced anomaly detection systems to strengthen security.

6. Conclusion

The progress of the 5G technology has induced many players in the enterprise as well as in the research fields to experiment various approaches in terms of scaling, performance improvement and technology availability. However, the field of security is still in early research and many features of the previous generation networks is being inherited. This entertains the disadvantage of attracting the very same vulnerabilities, which delved in those realms and lack the necessary means for approaching and resolving the same. One such example is the policy-based approach for enabling isolation and providing adjacent levels of security in 5G, but the same needs to be combined with other methods in order to be considered in more critical use-cases. The VPN technology has existed for a long time and has matured enough to be considered as a viable solution to privacy and security issues, nevertheless the same comes with several caveats. An example of such limitation is the hardware requirement for providing encrypted tunnels and needs a third-party solution in order to offload a generic machine for performing the dedicated tasks. Distribution at a scale from this aspect may not be appropriate for many actors, and thereby they will settle for ensuring their network slice isolation by other means.

Software-Defined Virtual Private Networks are an attractive approach, even though they incur a tremendous impact on the performance of a good link. For many use-cases, where the traffic needs an absolute protection and the particular network slice needs to be exclusively isolated (residing in a multitenancy environment, roaming or working with critical services), the VPN approach can be utilized outside the box and provide sufficient Quality of Service with the big advantage of an absolute privacy retention and unbreakable encryption. Methods for tackling the issues delivered by the encryption can be used, such as FPGA-based offloading or utilizing advanced traffic-steering techniques to instantiate multiple different tunnels, but these require more complex approach and further investigation. Additional research is needed to examine the potential benefit of combining the hardware-offloaded VPN tunneling and traffic-steering methods, as well as provide a more insightful approach to stimulate the Telco operators to consider the same at the transport network or even at the Edge.

This notion about employing Enhanced VPNs at the transport network can enable users to select ultra-secure network slices, with potentially same performance as a less secure high-bandwidth 5G slices if implemented in combination with techniques that will guarantee the

necessary Quality of Service. According to that, the users or vertical industries will not need to instantiate their private VPN applications and/or subscribe to third-parties for services which they do not ascertain the guaranteeing of their privacy. Providing a transport network encryption and tunneling on-demand can have benefits for the performances of remote network slices, which the users or vertical industries decide to utilize from another country and bypass the overlay encapsulation routing. These potential qualities are a solid starting point for further investigation and considering the experimented in this thesis. approach as a solid solution for many security issues that dwell into the unknown dominions of the incoming 5G world.

References

3GPP-TS 23.002. (2001). Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); Network architecture (3GPP TS 23.002 version 4.3.0 Release 4). Retrieved from https://www.etsi.org/deliver/etsi_ts/123000_123099/123002/04.03.00_60/ts_123002v040300p.pdf . Retrieved [Accessed Date March 2019], from etsi.org https://www.etsi.org/deliver/etsi_ts/123000_123099/123002/04.03.00_60/ts_123002v040300p.pdf

3GPP-TS 23.501. (2018). 5G - System Architecture for the 5G System (3GPP TS 23.501 version 15.2.0 Release 15) Retrieved from https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.02.00_60/ts_123501v150200p.pdf . Retrieved [Date Accessed Jan 2020], from etsi.org https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.02.00_60/ts_123501v150200p.pdf

3GPP-TS 23.501. (2019). 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; System architecture for the 5G System (5GS); Stage 2 (Release 16). Retrieved from https://www.3gpp.org/ftp/Specs/archive/23_series/23.501/. Retrieved [Date Accessed Dec 2019], from 3rd Generation Partnership Project (3GPP) https://www.3gpp.org/ftp/Specs/archive/23_series/23.501/

3GPP-TS 28.530. (2019). 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects;Management and orchestration;Concepts, use cases and requirements (Release 16). Retrieved from https://www.3gpp.org/ftp/Specs/archive/28_series/28.530/. Retrieved [Date Accessed Sep 2019], from 3rd Generation Partnership Project (3GPP) https://www.3gpp.org/ftp/Specs/archive/28_series/28.530/

3GPP-TS 33.102. (2012). 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security architecture (Release 15). Retrieved from http://www.3gpp.org/ftp/Specs/archive/33_series/33.102/. Retrieved [Date Accessed Dec 2019], from 3rd Generation Partnership Project 3GPP http://www.3gpp.org/ftp/Specs/archive/33_series/33.102/

3GPP-TS 33.103. (2001). 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G security Integration guidelines (Release 4). Retrieved from https://www.3gpp.org/ftp/tsg_sa/WG3_Security/_Specs/33103-420.pdf. Retrieved [Accessed Date Feb 2020], from 3rd Generation Partnership Project (3GPP) https://www.3gpp.org/ftp/tsg_sa/WG3_Security/_Specs/33103-420.pdf

3GPP-TS 37.470. (2020). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; W1 interface; General aspects and principles (Release 16). Retrieved from https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3468

3GPP-TS 38.211. (2019). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Physical channels and modulation (Release 15). Retrieved from https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213

3GPP-TS 38.401. (2020). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NG-RAN; Architecture description (Release 16). Retrieved from https://www.3gpp.org/ftp/Specs/archive/38_series/38.401/. Retrieved [Date Accessed Jan 2020], from 3rd Generation Partnership Project (3GPP) https://www.3gpp.org/ftp/Specs/archive/38_series/38.401/

3GPP-TS 38.473. (2018). 5G; NG-RAN; F1 Application Protocol (F1AP) (3GPP TS 38.473 version 15.2.1 Release 15) Retrieved from https://www.etsi.org/deliver/etsi_ts/138400_138499/138473/15.02.01_60/ts_138473v150201p.pdf . Retrieved [Date Accessed Feb 2020], from ETSI.org https://www.etsi.org/deliver/etsi_ts/138400_138499/138473/15.02.01_60/ts_138473v150201p.pdf

5GPPP. (2019). View on 5G Architecture (Whitepaper). Retrieved from https://5g-ppp.eu/wp-content/uploads/2019/07/5G-PPP-5G-Architecture-White-Paper_v3.0_PublicConsultation.pdf. Retrieved [Date Accessed July 2019], from 5GPPP https://5g-ppp.eu/wp-content/uploads/2019/07/5G-PPP-5G-Architecture-White-Paper_v3.0_PublicConsultation.pdf

Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., & Flinck, H. (2018). Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions. *IEEE Communications Surveys & Tutorials, 20*(3), 2429 - 2453. doi:10.1109/COMST.2018.2815638.

Aijaz, A. (2018). Hap-SliceR: A Radio Resource Slicing Framework for 5G Networks With Haptic Communications. *IEEE Systems Journal, 12*(3), 2285-2296. doi:10.1109/JSYST.2017.2647970

Ali, S., Faisal, E. M., Rahman, Z. U., & Haq, I. U. (2017). GSM Technology: Architecture, Security and Future Challenges. *International Journal of Science Engineering and Advance Technology IJSEAT, 5*(1), 70-74. Retrieved from http://www.ijseat.com/index.php/ijseat/article/view/789/pdf

Anderson, R. J. (2008). *Security Engineering: A Guide to Building Dependable Distributed Systems* (Second ed.). Indianapolis, USA: Wiley Publishing Inc.

Bais, A., T. Penzhorn, W., & Palensky, P. (2006). *Evaluation of UMTS security architecture and services*. Paper presented at the 2006 4th IEEE International Conference on Industrial Informatics, Singapore, Singapore.

Barakabitze Alcardo, A., Ahmad, A., Mijumbi, R., & Hines, A. (2020). 5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges. *Computer Networks, 167*(Special Selection). doi:https://doi.org/10.1016/j.comnet.2019.106984

Bauer, M., Schefczik, P., Soellner, M., & Speltacker, W. (2003). *Evolution of the UTRAN Architecture*. Paper presented at the 4th International Conference on 3G Mobile Communication Technologies, London, UK, UK.

Bertenyi, B., Burbidge, R., Masini, G., Sirotkin, S., & Gao, Y. (2018). NG Radio Access Network (NG-RAN). *Journal of ICT Standardization, 6*, 59-76 doi:https://doi.org/10.13052/jicts2245-800X.614

Brydon, A. (2013, December, 2019). 4x2 MIMO (Multiple Input Multiple Output). Retrieved from https://www.unwiredinsight.com/2013/lte-mimo

Brzozowski, A. (2020). Norway takes 'most far-reaching measures ever experienced in peacetime', plans easing in April. *Coronavirus.* Retrieved from https://www.euractiv.com/section/coronavirus/short_news/norway-update-covid-19/

Cattaneo, G., De Maio, G., & Petrillo, U. (2013). Security Issues and Attacks on the GSM Standard: a Review. *Journal of Universal Computer Science, 19*, 2437-2452. doi:10.3217/jucs-019-16-2437

Ceccarelli, D., & Lee, Y. (2018). Framework for Abstraction and Control of Traffic Engineered Networks (ACTN). RFC 8453. Retrieved from https://tools.ietf.org/html/draft-ietf-teas-actn-framework-15

Chandra, P. (2005). *Bulletproof Wireless Security: GSM, UMTS, 802.11 and Ad Hoc Security*: Elsevier.

Chayapathi, R., Hassan, S. F., & Shah, P. (2016). *Network Functions Virtualization (NFV) with a Touch of SDN* (First ed.): Addison-Wesley Professional.

Chen, J. C., & Zhang, T. (2004). *Introduction IP-Based Next-Generation Wireless Networks Systems, Architectures, and Protocols* John Wiley & Sons, Inc.

Chouchane, A., Rekhis, S., & Boudriga, N. (2009). *Defending against rogue base station attacks using wavelet based fingerprinting*. Paper presented at the 2009 IEEE/ACS International Conference on Computer Systems and Applications, Rabat, Morocco. https://ieeexplore.ieee.org/document/5069374

Cisco. (2006). Point-to-Point GRE over IPsec Design Guide. Retrieved from Available from World Wide Web: <https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/WAN_and_MAN/P2P_GRE_IPSec/

P2P_GRE/2_p2pGRE_Phase2.html>. Retrieved May, 2020, from Cisco Systems Available from World Wide Web: <https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/WAN_and_MAN/P2P_GRE_IPSec/P2P_GRE/2_p2pGRE_Phase2.html>

Cisco. (2020). Cisco Annual Internet Report (2018–2023) White Paper. Retrieved from [Available from World Wide Web: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html]

Das, S. K. (2010). *Mobile Handset Design*. Wiley online library: John Wiley & Sons (Asia) Pte Ltd. .

Dell Inc. (2012). Dell PowerEdge R620 (Technical Specification). Retrieved from https://www.dell.com/downloads/global/products/pedge/en/Dell_PowerEdge_R620_Spec_Sheet.pdf. Retrieved January, 2020, from Dell https://www.dell.com/downloads/global/products/pedge/en/Dell_PowerEdge_R620_Spec_Sheet.pdf

Devra, S., & Sharma, A. (2016). Second International Conference On Innovative Trends In Electronics Engineering (ICITEE2) GSM Architecture & Channels: Review Study. Retrieved from http://ijoes.vidyapublications.com/paper/Vol20/02-Vol20.pdf. Retrieved June, 2020, from Vidya Publications http://ijoes.vidyapublications.com/paper/Vol20/02-Vol20.pdf

Dhawan, S. K. (2005). *Introduction to PCI Express-a new high speed serial data bus*. Paper presented at the IEEE Nuclear Science Symposium Conference Record, 2005, Fajardo, Puerto Rico.

Docker. (2020). Docker SR-IOV plugin. Retrieved from https://hub.docker.com/r/rdma/sriov-plugin

Docker Inc. (2020a). Developers bring their ideas to life with Docker. Retrieved from https://www.docker.com/why-docker

Docker Inc. (2020b). What is a Container? Retrieved from https://www.docker.com/resources/what-container

Doll, M., Sciancalepore, V., Bega, D., Schneider, P., Rost, P., & S, M. D. (2017). *Network slicing via function decomposition and flexible network design*. Paper presented at the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada.

Dong, J., Li, Z., Miyasaka, T., Bryant, S., & Young, L. (2020). A Framework for Enhanced Virtual Private Networks (VPN+) Services. Retrieved from https://datatracker.ietf.org/doc/draft-ietf-teas-enhanced-vpn/

Dong, Y., Yang, X., Li, X., Li, J., Tian, K., & Guan, H. (2010). *High performance network virtualization with SR-IOV*. Paper presented at the The Sixteenth International Symposium on High-Performance Computer Architecture (HPCA), Bangalore, India.

Dzogovic, B., Do, v. T., Feng, B., & Do, v. T. (2018). *Building virtualized 5G networks using open source software*. Paper presented at the 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), Penang, Malaysya.

Dzogovic, B., Do, v. T., Santos, B., Do, v. T., Feng, B., & Jacot, N. (2019). *Thunderbolt-3 Backbone for Augmented 5G Network Slicing in Cloud-Radio Access Networks*. Paper presented at the 2019 IEEE 2nd 5G World Forum (5GWF), Dresden, Germany.

Dzogovic, B., Do, V. T., Satnos, B., Do, T. V., Feng, B., & Jacot, N. (2019). *Connecting Remote eNodeB with Containerized 5G C-RANs in OpenStack Cloud*. Paper presented at the 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), Paris, France. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8854015>

Eberspacher, J., Vogel, H. J., & Bettstetter, C. (2001). *GSM Switching, Services and Protocols* (Second ed.): John Wiley & Sons, Ltd.

Elvidge, A. M., & Martucci, J. (2003). Telecommunications Network Total Cost of Ownership and Return on Investment Modelling. *BT Technology Journal by Springer Link, 21*, 184–190. doi:https://doi.org/10.1023/A:1024467706446

Ericsson. (2018). Network Slicing can be a piece of cake. Retrieved from [Available from World Wide Web:< https://www.ericsson.com/assets/local/digital-services/doc/Network-slicing-Report-Final.pdf?_ga=2.152704185.1839582109.1589384116-430220867.1580331862>]. Retrieved [Date Accessed March 2020], from Ericcson Press release and paper [Available from World Wide Web:< https://www.ericsson.com/assets/local/digital-services/doc/Network-slicing-Report-Final.pdf?_ga=2.152704185.1839582109.1589384116-430220867.1580331862>]

Ericsson. (2020). Ericsson Spectrum Sharing for immersive, wide-area 5G coverage. Retrieved from [Available from World Wide Web:< https://www.ericsson.com/en/networks/offerings/5g/sharing-spectrum-with-ericsson-spectrum-sharing>]

ETSI-GSM 02.16. (2000). Digital cellular telecommunications system (Phase 2); International Mobile station Equipment Identities (IMEI) (GSM 02.16 version 4.7.1).[Online]. Fourth. Retrieved from https://www.etsi.org/deliver/etsi_i_ets/300500_300599/300508/04_60/ets_300508e04p.pdf

ETSI-GSM 05.01. (1992). European Digital Cellular Telecommunications System (phase 1); Physical Layer on the Radio Path General description. V3.2.2. Retrieved from https://www.etsi.org/deliver/etsi_gts/05/0501/03.03.02_60/gsmts_0501sv030302p.pdf

ETSI-GSM 08.52. (1996). Digital cellular telecommunications system; Base Station Controller - Base Transceiver Station (BSC - BTS) interface; Interface principles (GSM 08.52). Retrieved from https://www.etsi.org/deliver/etsi_gts/08/0852/05.00.00_60/gsmts_0852v050000p.pdf

ETSI-TS 23.119. (2012). Universal Mobile Telecommunications System (UMTS); LTE; Gateway Location Register (GLR); Stage2 (3GPP TS 23.119 version 11.0.0 Release 11) Retrieved from https://www.etsi.org/deliver/etsi_ts/123100_123199/123119/11.00.00_60/ts_123119v110000p.pdf

ETSI-TS 23.236. (2004). Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Intra-domain connection of Radio Access Network (RAN) nodes to multiple Core Network (CN) nodes (3GPP TS 23.236 version 6.0.0 Release 6). Retrieved from https://www.etsi.org/deliver/etsi_ts/123200_123299/123236/06.00.00_60/ts_123236v060000p.pdf

ETSI-TS 36.304. (2014). LTE;Evolved Universal Terrestrial Radio Access (E-UTRA);User Equipment (UE) procedures in idle mode (3GPP TS 36.304 version 11.6.0 Release 11) Retrieved from https://www.etsi.org/deliver/etsi_ts/136300_136399/136304/11.06.00_60/ts_136304v110600p.pdf

Ettus Research, (2020). Universal Software Radio Peripheral (USRP) B210. Retrieved from https://www.ettus.com/all-products/ub210-kit/

EURECOM. (2020). OpenAirInterface5G software alliance for democratising wireless innovation Retrieved from https://www.openairinterface.org/

F.Naranjo, E., & D.Salazar Ch, G. (2017). *Underlay and overlay networks: The approach to solve addressing and segmentation problems in the new networking era: VXLAN encapsulation with Cisco and open source networks*. Paper presented at the 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), Salinas, Ecuador. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8247505>

Farrel, A. (2017). TO SLICE OR NOT TO SLICE, THAT IS THE QUESTION. Retrieved from Available from World Wide Web: <https://metro-haul.eu/2017/11/08/to-slice-or-not-to-slice-that-is-the-question/>

Farrel, A. (2018). WHAT IS ACTN? Retrieved from https://metro-haul.eu/2018/08/30virtu/what-is-actn/

Farrel, A. (2019a). EVOLVING VPN-BASED SERVICES: A FRAMEWORK FOR ENHANCED VPNS. In: Metro-Haul.

Farrel, A. (2019b). Network Slicing and Enhanced VPNs. Retrieved from [Available from World Wide Web:< http://www.olddog.co.uk/Farrel-VPN.pdf>]

Farrel, A., Drake , J., Bitar, N., Swallow, G., Ceccarelli, D., & Zhang, X. (2016, [Date Accessed March 2020]). Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks Retrieved from [Available from World Wide Web:< https://rfc-editor.org/rfc/rfc7926.txt>]

Feng, B., Do, T. V., Jacot, N., Santos, B., Dzogovic, B., Brandsma, E., & Do, V. T. (2019). *Secure 5G Network Slicing for Elderly Care*. Paper presented at the 16th International Conference on Mobile Web and Intelligent Information Systems (MobiWIS 2019), Istanbul, Turkey.

Firmin, F., & 3GPP, M. (2020). The Evolved Packet Core. Retrieved from [Available from World Wide Web:< https://www.3gpp.org/technologies/keywords-acronyms/100-the-evolved-packet-core>]

Foukas, X., Patounas, G., Elmokashfi, A., & Marina, M. K. (2017). Network Slicing in 5G: Survey and Challenges. *IEEE Communications Magazine, 55(5),* 94-100. Retrieved from Available from World Wide Web:< http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7926923&isnumber=7926739>

Garg, V. K., & Wilkes, J. E. (1998). *Principles and Applications of GSM*. Upper Saddle River, NJ United States: Prentice Hall PTR.

Gont, F. (2014). Layer 3 Virtual Private Network (VPN) Tunnel Traffic Leakages in Dual-Stack Hosts/Networks. Internet Engineering Task Force (IETF) RFC 7359. Retrieved from [Availble from World Wide Web:< https://www.rfc-editor.org/info/rfc7359>]

Gsma. (2018). Guidelines for IPX Provider networks (Previously InterService Provider IP Backbone Guidelines) Version 14.0. Retrieved from [Available from World Wide Web:< https://www.gsma.com/newsroom/wp-content/uploads//IR.34-v14.0.pdf>]. Retrieved [Date Accessed March 2020], from gsma [Available from World Wide Web:< https://www.gsma.com/newsroom/wp-content/uploads//IR.34-v14.0.pdf>]

Gsma. (2019a). THE 5G GUIDE A REFERENCE FOR OPERATORS. Retrieved from Avaibale from World Wide Web: < https://www.gsma.com/wp-content/uploads/2019/04/The-5G-Guide_GSMA_2019_04_29_compressed.pdf>. Retrieved [Date Accessed May 2020], from gsma Avaibale from World Wide Web: < https://www.gsma.com/wp-content/uploads/2019/04/The-5G-Guide_GSMA_2019_04_29_compressed.pdf>

Gsma. (2019b). Brief History of GSM & the GSMA. Retrieved from [Available from World Wide web: <https://www.gsma.com/aboutus/history>]

Gsma. (2019c). Representing the worldwide mobile communications industry. Retrieved from [Available from World Wide Wibe:< https://www.gsma.com/>]

Gsma. (2019d). Security Algorithms. *security.* Retrieved from [Available from World Wide Web:< https://www.gsma.com/security/security-algorithms/>]

gsmareana. (2019). Retrieved from https://www.gsmarena.com/t_mobile_mytouch_3g-pictures-2848.php

Harri, H., & Antti, T. (2009). *LTE for UMTS: OFDMA and SC-FDMA based radio access* (1 ed.): John Wiley & Sons.

Harte, L., Bromley, B., & Davis, M. (2008). Introduction to GSM: Physical Channels, Logical Channels, Network Functions, and Operation. In: Phoenix Global Support.

Hartung, F., Horn, U., Kampmann, M., Lohmar, T., & Huschke, J. (2009). MBMS—IP Multicast/Broadcast in 3G Networks. *International Journal of Digital and Multimedia Broadcasting, 2009*, 25. doi:https://doi.org/10.1155/2009/597848

He, L., Yan, Z., & Atiquzzaman, M. (2018). LTE/LTE-A Network Security Data Collection and Analysis for Security Measurement: A Survey. *IEEE Access, 6*, 4220 - 4242. doi:10.1109/ACCESS.2018.2792534

Heine, G. (1999). *GSM Networks: Protocols, Terminology, and Implementation*: Artech House Publishers.

Henrydoss, J., & Boult, T. (2014). *Critical security review and study of DDoS attacks on LTE mobile network*. Paper presented at the 2014 IEEE Asia Pacific Conference on Wireless and Mobile, Bali, Indonesia. Available from World Wide Web: <https://ieeexplore.ieee.org/document/6920286>

Hirayama, H., Tsukamoto, Y., Nanba, S., & Nishimura, K. (2019). *RAN Slicing in Multi-CU/DU archiecture for 5G Services*. Paper presented at the 2019 IEEE 90th Vehicular Techonology Conference (VTC2019-Fall), Honolulu, HI, USA, USA. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8891584>

Idrissi, D. E., Elkamoun, N., & Hilal, R. (2019). *Study of the impact of failure on GRE Tunnel*. Paper presented at the 2019 7th Mediterranean Congress of Telecommunications (CMT), Fès, Morocco, Morocco. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8931368>

IETF. (2009). RFC5440: Path Computation Element (PCE) Communication Protocol (PCEP). Retrieved from https://tools.ietf.org/html/rfc5440. from Internet Engineering Task Force (IETF) https://tools.ietf.org/html/rfc5440

IETF. (2015). RFC7540: Hypertext Transfer Protocol Version 2 (HTTP/2). Retrieved from https://www.rfc-editor.org/info/rfc7540. Retrieved [Date Accessed Feb 2020], from Internet Engineering Task Force (IETF) https://www.rfc-editor.org/info/rfc7540

IETF. (2016). RFC7926: Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks. In: IETF.

Ijaz, A., Tanesh, K., Madhusanka, L., Jude, O., Mika, Y., & Andrei, G. (2018). Overview of 5G Security Challenges and Solutions. *IEEE Communications Standards Magazine 2(1),* 36-43. Retrieved from Available from World Wide Web: <https://ieeexplore.ieee.org/document/8334918>

Intel. (2011). PCI-SIG SR-IOV Primer An Introduction to SR-IOV Technology. Retrieved from Available from World Wide Web:< https://www.intel.com/content/www/us/en/pci-express/pci-sig-sr-iov-primer-sr-iov-technology-paper.html>. Retrieved [Accessed Date March 2020], from Intel® LAN Access Division Available from World Wide Web:< https://www.intel.com/content/www/us/en/pci-express/pci-sig-sr-iov-primer-sr-iov-technology-paper.html>

Intel. (2017). Enabling New Features with Kubernetes for NFV (whitepaper), Intel Corporation Data Center Solution Networking Group. Retrieved from Avialable from World Wide Web: <https://builders.intel.com/docs/networkbuilders/enabling_new_features_in_kubernetes_for_NFV.pdf>

Interdynamix Systems. (2018). 5G NETWORK SLICING AND OPENSTACK. In: idxLABS.

iPerf3. (2020). iPerf3 testing tool Retrieved from Available from World Wide Web:< https://iperf.fr/>

ITU. (2018). *Transport network support of IMT-2020/5G*. Retrieved from Ciena, Canada: https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-HOME-2018-PDF-E.pdf

Jain, V., Singh, T., & Babu, G. S. (2017). *VXLAN and EVPN for data center network transformation*. Paper presented at the 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Delhi, India. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8203947>

Juniper Networks. (2011). Overview of PDP Contexts and Bearers. Retrieved from Available from World Wide Web: <https://www.juniper.net/documentation/en_US/junos-mobility11.2/topics/concept/gateways-mobility-bearer-overview.html>

Juniper Networks. (2015). LTE Security for Mobile Service Provider Networks (White paper). Retrieved from Available from World Wide Web: < https://www.juniper.net/us/en/local/pdf/whitepapers/2000536-en.pdf>. Retrieved [Accessed Date May 2020], from Juniper Networks Available from World Wide Web: < https://www.juniper.net/us/en/local/pdf/whitepapers/2000536-en.pdf>

Jyothi, K. K., & Reddy, D. B. I. (2018). Study on Virtual Private Network (VPN), VPN's Protocols And Security. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 3*(5), 919-932. Retrieved from Available from World Wide Web: <http://ijsrcseit.com/CSEIT1835225>

Kaaranen, H., Ahtiainen, A., Laitinen, L., Naghian, S., & Niemi, V. (2005). *UMTS networks: architecture, mobility and services* (Second ed.): John Wiley & Sons.

Kapassa, E., Touloupou, M., Mavrogiorgou, A., Kiourtis, A., Giannouli, D., Katsigianni, K., & Kyriazis, D. (2019). *An Innovative eHealth System Powered By 5G Network Slicing*. Paper presented at the 2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS), Granada, Spain, Spain. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8939266>

Karim, M., & Sarraf, M. (2002). *W-CDMA and cdma2000 for 3G Mobile Networks* (First ed.): McGraw-Hill Professional.

Kazmi, S. M. A., Khan, L. U., H.Tran, N., & Hong, C. S. (2019). *Network Slicing for 5G and Beyond Networks* (First ed.): Springer.

Khan, M. Q. (2018). *Signaling Storm Problems in 3GPP Mobile Broadband Networks, Causes and Possible Solutions: A Review*. Paper presented at the 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), Southend, United Kingdom, United Kingdom. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8658708>

Khettab, Y., Bagaa, M., Cadette Dutra, D. L., Taleb, T., & Toumi, N. (2018). *Virtual security as a service for 5G verticals*. Paper presented at the 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain. Available from World Wide Web: <https://ieeexplore.ieee.org/document/8377298/>

Kokku, R., Mahindra, R., Zhang, H., & Rangarajan, S. (2013). *CellSlice: Cellular wireless resource slicing for active RAN sharing*. Paper presented at the 2013 Fifth International Conference on Communication Systems and Networks (COMSNETS), Bangalore, India. Available from World Wide Web: <https://ieeexplore.ieee.org/document/6465548/>

Korhonen, J. (2003). *Introduction to 3G Mobile Communications* (Second ed.): Artech House.

Kumar, A., Aswal, A., & Singh, L. (2013). 4G Wireless Technology: A Brief Review. *International Journal of Engineering and Management Research, 3*(2), 35-43. Retrieved from Available from World Wide Web: < http://www.ijemr.net/DOC/4GWirelessTechnology-%20ABriefReview(35-43)50dde7d5-1653-4f14-b384-5da6783e0ffe.pdf>

Kumar, D. S., & Meraj ud in Mir, M. (2015). Evolution of mobile wireless technology from 0G to 5G. *International Journal of Computer Science and Information Technologies (IJCSIT), 6*(3), 2545-2551. Retrieved from Available from World Wide Web <https://ijcsit.com/docs/Volume%206/vol6issue03/ijcsit20150603123.pdf>

Kussum, B., & Malhotra, J. (2014). *A survey of uplink multiple access techniques in LTE mobile communication system*. Paper presented at the 2014 International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), Unnao, India. Available from World Wide Web: <https://ieeexplore.ieee.org/document/7012909>

Lal, S., Taleb, T., & Dutta, A. (2017). NFV: Security Threats and Best Practices. *IEEE Communications Magazine, 55(8),* 211-217. Retrieved from Available from World Wide Web: <https://ieeexplore.ieee.org/document/7927935>

Lee, J., Han, J.-K., & Jianzhong, Z. (2009). MIMO technologies in 3GPP LTE and LTE-advanced. *EURASIP Journal on Wireless Communications and Networking*. doi:https://doi.org/10.1155/2009/302092

Lee, Y., & Kaippallimalil, J. (2019). Applicability of ACTN to Support 5G Transport. Retrieved from Available from World Wide Web: <https://datatracker.ietf.org/doc/draft-lee-teas-actn-5g-transport/>. from Internet Engineering Task Force (IETF) Available from World Wide Web: <https://datatracker.ietf.org/doc/draft-lee-teas-actn-5g-transport/>

Linge, N. (2019). Engaging with communications. Retrieved from http://www.engagingwithcommunications.com/Technology/Mobiles/TACS/tacs.php

Liyanage, M., Ahmad, I., Abro, A. B., Gurtov, A., & Ylianttila, M. (2018). *A Comprehensive Guide to 5G Security*: John Wiley & Sons.

Liyanage, M., & Gurtov, A. (2012). *Secured VPN Models for LTE Backhaul Networks*. Paper presented at the 2012 IEEE Vehicular Technology Conference (VTC Fall), Quebec City, QC, Canada. Available from World Wide Web: <https://ieeexplore.ieee.org/document/6399037>

Liyanage, M., Kumar, P., Ylianttila, M., & Gurtov, A. (2016). Novel secure VPN architectures for LTE backhaul networks (Publication no. 10.1002/sec.1411). Retrieved [Date Accessed April 2020], from John Wiley & Sons, Ltd. Available from World Wide Web: <https://onlinelibrary.wiley.com/doi/full/10.1002/sec.1411>

Masrom, M., & Ali, A. h. (2010). *Analysis and implementation of security algorithms for wireless communications*. Paper presented at the 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore, Singapore. Available from World Wide Web: <https://ieeexplore.ieee.org/document/5451589>

May, D. (2013). Function usage provides the ability for more granular security controls. Retrieved from Available from World Wide Web:< https://www.ibm.com/developerworks/ibmi/library/i-granular-security/index.html>. Retrieved [Accessed Date Nov 2019], from IBM Available from World Wide Web:< https://www.ibm.com/developerworks/ibmi/library/i-granular-security/index.html>

McMillan, R., Knutson, R., & Seetharaman, D. (2016). Yahoo Discloses New Breach of 1 Billion User Accounts. Retrieved from https://www.wsj.com/articles/yahoo-discloses-new-breach-of-1-billion-user-accounts-1481753131

Medium. (2018). 5G Service-Based Architecture (SBA). Retrieved from https://medium.com/5g-nr/5g-service-based-architecture-sba-47900b0ded0a

Mellanox Technologies. (2020). Docker SR-IOV passthrough plugin Retrieved from https://github.com/Mellanox/docker-sriov-plugin

Mosaic-5G. (2020). FlexRAN controller. Retrieved from http://mosaic-5g.io/flexran/

NGMN. (2015). NGMN 5G White paper. Retrieved from Available from World Wide Web: <https://www.ngmn.org/wp-content/uploads/NGMN_5G_White_Paper_V1_0.pdf>. Retrieved [Date Accessed Dec 2019], from NGMN Alliance Available from World Wide Web: <https://www.ngmn.org/wp-content/uploads/NGMN_5G_White_Paper_V1_0.pdf>

NGMN. (2019). 5G RAN CU - DU Network Architecture, Transport Options and Dimensioning v 1.0 Retrieved from Available from World Wide Web:< https://www.ngmn.org/wp-content/uploads/Publications/2019/190412_NGMN_RANFSX_D2a_v1.0.pdf>. Retrieved [Accessed date Nov 2019], from NGMN Alliance Available from World Wide Web:< https://www.ngmn.org/wp-content/uploads/Publications/2019/190412_NGMN_RANFSX_D2a_v1.0.pdf>

Nikaein, N., Marina, M. K., Manickam, S., Dawson, A., Knopp, R., & Bonnet, C. (2014). OpenAirInterface: A Flexible Platform for 5G Research. *ACM SIGCOMM Computer Communication Review, 44*(5). Retrieved from Available from World Wide Web: <https://doi.org/10.1145/2677046.2677053>

Nohrborg, M. (2019). LTE overveiw 3GPP. Retrieved from Available from World Wide Web: <https://www.3gpp.org/technologies/keywords-acronyms/98-lte>

Nokia Networks Oy. (2001). GSM Architecture Training Document, . Retrieved from https://www.academia.edu/4221298/GSM_Architecture_Training_Document_TC_Finland. Retrieved March, 2019, from Nokia Networks https://www.academia.edu/4221298/GSM_Architecture_Training_Document_TC_Finland

Olsson, M., Mulligan, C., Sultana, S., Rommer, S., & Frid, L. (2012). *EPC and 4G packet networks: driving the mobile broadband revolution* (Second ed.): Academic Press.

Onos Projects. (2018). ONOS New Projects "ACTN (Abstraction and Control of TE networks)". Retrieved from Available from World Wide Web: <https://wiki.onosproject.org/pages/viewpage.action?pageId=8424694>

OpenStack. (2020a). Kuryr plugin for container networking. Retrieved from https://wiki.openstack.org/wiki/Kuryr

OpenStack. (2020b). Neutron networking service documentation. Retrieved from https://docs.openstack.org/neutron/pike/index.html

Ordonez-Lucena, J., Amerigeiras, P., Lopez, D., J.Ramos-Munoz, J., Lorca, J., & Folgueira, J. (2017). Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges. *IEEE Communications Magazine, 55(5),* 80-87. Retrieved from https://ieeexplore.ieee.org/document/7926921

Palanivelu, T. G., & Nakkeeran, R. (2008). *Wireless and Mobile Communication*: PHI Learning Private Limited.
Rahnema, M. (1993). Overview of the GSM system and protocol architecture. *IEEE Communications Magazine, 31(4),* 92-100. Retrieved from https://ieeexplore.ieee.org/document/210402
Redhat. (2020). VIRTUALIZATION ,What is NFV? Retrieved from https://www.redhat.com/en/topics/virtualization/what-is-nfv
RF Wireless World. (2012). Home of RF and Wireless Vendors and Resources One Stop For Your RF and Wireless Need. Retrieved from https://www.rfwireless-world.com/Tutorials/gsm-channel-types.html
Rocha, F., & Correia, M. (2011). *Lucy in the sky without diamonds: Stealing confidential data in the cloud*. Paper presented at the 2011 IEEE/IFIP 41st International Conference on Dependable Systems and Networks Workshops (DSN-W), Hong Kong.
S. Kanchi, S. S., D. Bhosale, A. Pitkar and M. Gondhalekar;. (2013). "Overview of LTE-A technology," 2013 IEEE Global High Tech Congress on Electronics, Shenzhen, pp. 195-200. doi: 10.1109/GHTCE.2013.6767272. Retrieved from [Document link http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6767272&isnumber=6767224]
Samdanis, K., Wright, S., Banchs, A., Capone, A., Ulema, M., & Obana, K. (2017). 5G Network Slicing: Part 1 – Concepts, Principles, and Architectures. *IEEE Communications Magazine, 55(5),* 70-71. Retrieved from https://ieeexplore.ieee.org/document/7926919
Schmitt, P., Landais, B., & Yang, F. Y. (2017). Control and User Plane Separation of EPC nodes (CUPS). Retrieved from https://www.3gpp.org/news-events/1882-cups
Sempere, J. G. (2002). An overview of the GSM system. Retrieved from https://web.fe.up.pt/~mleitao/CMOV/Tecnico/GSM_Sempere.html
Shaik, A., Borgaonkar, R., Asokan, N., Niemi, V., & Seifert, J.-P. (2016). *Practical attacks against privacy and availability in 4G/LTE mobile communication systems*. Paper presented at the 23rd Annual Network and Distributed System Security Symposium (NDSS 2016) Internet Society. https://arxiv.org/pdf/1510.07563.pdf
Singh, J., Ruhl, R., & Lindskog, D. (2013). *GSM OTA SIM Cloning Attack and Cloning Resistance in EAP-SIM and USIM*. Paper presented at the 2013 International Conference on Social Computing, Alexandria, VA, USA.
Snader, J. C. (2005). *VPNs Illustrated: Tunnels, VPNs, and IPsec* (1st ed.): Addison-Wesley Professional.
Stackoverflow. (2020). Jason Questions And Answers. Retrieved from https://stackoverflow.com/questions/tagged/json
Taleb, T., Ksentini, A., & Sericola, B. (2016). On Service Resilience in Cloud-Native 5G Mobile Systems. *IEEE Journal on Selected Areas in Communications, 34*(3), 483-496. doi:10.1109/JSAC.2016.2525342
Terasvuo, K. (2019). History of GSM - Birth of the Mobile Revolution. Retrieved from http://www.gsmhistory.com/the-beginnings/.
Thampi, S. M., Madria, S., Wang, G., Rawat, D. B., & Calero, J. M. A. (2019). *Security in Computing and Communications* (Vol. 969). Bangalore, India: Springer.
Vohra, D., Dubey, A., & Vachhhani, K. (2016). *Investigating GSM Control Channels with RTL-SDR and GNU Radio*. Paper presented at the 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India. https://ieeexplore.ieee.org/document/7566288
Wang, Y. (2015). *Network virtualization over SLICE networks*. Paper presented at the 2015 36th IEEE Sarnoff Symposium, Newark, NJ, USA. https://ieeexplore.ieee.org/document/7324659
Wolff, J. (2020). Our Internet Isn't Ready for Coronavirus by Dr. Wolff is an assistant professor at Tufts University and a contributing opinion writer. Retrieved from https://www.nytimes.com/2020/03/17/opinion/coronavirus-broadband-internet-work-from-home.html

Xiang, W., Zheng, K., & Shen, X. S. (2017). *5G mobile communications*. Online: Springer Publishing International.
Yang, W., & Fung, C. (2016). *A survey on security in network functions virtualization*. Paper presented at the 2016 IEEE NetSoft Conference and Workshops (NetSoft), Seoul, South Korea.
Zhang, Y. (2018). *Network Function Virtualization: Concepts and Applicability in 5G Networks* (First ed.): Wiley-IEEE Press.

Appendix

A) 5G Core host networking configuration. The container has a different MAC address from the eno2 interface to which the SR-IOV virtualization is mapped. Nevertheless, the container is considered as a member of the same network

```
eno1: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 9000
        inet 128.39.120.239  netmask 255.255.254.0  broadcast 128.39.121.255
        inet6 fe80::d294:66ff:fe09:4f2f  prefixlen 64  scopeid 0x20<link>
        inet6 2001:700:700:22:d294:66ff:fe09:4f2f  prefixlen 64  scopeid 0x0<global>
        ether d0:94:66:09:4f:2f  txqueuelen 1000  (Ethernet)
        RX packets 90221191  bytes 46188378351 (46.1 GB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 86420819  bytes 47004318688 (47.0 GB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 68  memory 0x95000000-957fffff

eno2: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 9000
        inet 10.0.0.1  netmask 255.255.255.0  broadcast 10.0.0.255
        inet6 fe80::d294:66ff:fe09:4f31  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:09:4f:31  txqueuelen 1000  (Ethernet)
        RX packets 123284167  bytes 1073889130460 (1.0 TB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 55843942  bytes 78402855764 (78.4 GB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x94000000-947fffff

eno3: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        ether d0:94:66:09:4f:33  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x93000000-937fffff

eno4: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 9000
        inet 192.168.1.1  netmask 255.255.255.0  broadcast 192.168.1.255
        inet6 fe80::d294:66ff:fe09:4f35  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:09:4f:35  txqueuelen 1000  (Ethernet)
        RX packets 1026564  bytes 87978093 (87.9 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 44658  bytes 2210592 (2.2 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 103  memory 0x92000000-927fffff

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

root@atlas:~# docker exec -it oai_epc /bin/bash
root@epc:/# ifconfig
eth0      Link encap:Ethernet  HWaddr 02:42:0a:00:00:03
          inet addr:10.0.0.3  Bcast:10.0.0.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:9000  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:65536  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)
```

B) Centralized Unit host networking configuration

```
root@poaseidon:~# ifconfig
docker0: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        inet 172.17.0.1  netmask 255.255.0.0  broadcast 172.17.255.255
        ether 02:42:82:ad:65:19  txqueuelen 0  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

eno1: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 9000
        inet 10.0.0.2  netmask 255.255.255.0  broadcast 10.0.0.255
        inet6 fe80::d294:66ff:fe02:17b1  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:02:17:b1  txqueuelen 1000  (Ethernet)
        RX packets 55852656  bytes 78403813609 (78.4 GB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 123278818  bytes 1073888507370 (1.0 TB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 68  memory 0x95000000-957fffff

eno2: flags=4099<UP,BROADCAST,MULTICAST>  mtu 9000
        ether d0:94:66:02:17:b3  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x94000000-947fffff

eno3: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        ether d0:94:66:02:17:b5  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x93000000-937fffff

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 204169  bytes 17173710 (17.1 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 204169  bytes 17173710 (17.1 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

root@poaseidon:~# ip r
default via 10.0.0.1 dev eno1 proto static
10.0.0.0/24 dev eno1 proto kernel scope link src 10.0.0.2
172.17.0.0/16 dev docker0 proto kernel scope link src 172.17.0.1 linkdown
root@poaseidon:~#
```

C) 5G Core Network AMF/MME function listening for Centralized Unit Connection (registered as an eNB/gNB). Parent interfaces are formed and bridged to the *eth0*, which is a SR-IOV passthrough (with its own MAC address and IP=10.0.0.3/24, making the container part of a local network as it is a physical machine.

```
00510 00270:757686 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0034                 |   Current Status| Added since last d
isplay|  Removed since last display |
00511 00270:757695 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0036    Connected eNBs |          0      |             0
         |             0               |
00512 00270:757703 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0038    Attached UEs   |          0      |             0
         |             0               |
00513 00270:757711 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0040    Connected UEs  |          0      |             0
         |             0               |
00514 00270:757722 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0042    Default Bearers|          0      |             0
         |             0               |
00515 00270:757732 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0044    S1-U Bearers   |          0      |             0
         |             0               |

00516 00270:757742 7F6A21408700 DEBUG MME-AP src/mme_app/mme_app_statistics.c:0045    ==================================== STATISTICS ==
========================================


root@epc:/usr/local/etc/oai# ifconfig
eth0      Link encap:Ethernet  HWaddr 02:42:0a:00:00:03
          inet addr:10.0.0.3  Bcast:10.0.0.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:9000  Metric:1
          RX packets:131 errors:0 dropped:0 overruns:0 frame:0
          TX packets:22 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:38394 (38.3 KB)  TX bytes:1428 (1.4 KB)

eth0:11   Link encap:Ethernet  HWaddr 02:42:0a:00:00:03
          inet addr:192.171.11.1  Bcast:192.171.11.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:9000  Metric:1

eth0:21   Link encap:Ethernet  HWaddr 02:42:0a:00:00:03
          inet addr:192.171.11.2  Bcast:192.171.11.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:9000  Metric:1

gtp0      Link encap:UNSPEC  HWaddr 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00
          inet addr:192.188.0.1  P-t-P:192.188.0.1  Mask:255.255.255.0
          UP POINTOPOINT RUNNING NOARP MULTICAST  MTU:9000  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          UP LOOPBACK RUNNING  MTU:65536  Metric:1
          RX packets:165 errors:0 dropped:0 overruns:0 frame:0
          TX packets:165 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:22274 (22.2 KB)  TX bytes:22274 (22.2 KB)

root@epc:/usr/local/etc/oai#
```

D) Centralized Unit configuration file. The node servers the purpose of a IF4p5 (Option-7) functional split for Band7, Sub-6 spectrum (2685 MHz).

```
Active_eNBs = ( "eNB-Eurecom-LTEBox");
# Asn1_verbosity, choice in: none, info, annoying
Asn1_verbosity = "none";

eNBs =
(
 {
    # real_time choice in {hard, rt-preempt, no}
    real_time       =  "no";

    ////////// Identification parameters:
    eNB_ID    =  0xe00;

    cell_type =  "CELL_MACRO_ENB";

    eNB_name  =  "eNB-Eurecom-LTEBox";

    // Tracking area code, 0x0000 and 0xfffe are reserved values
    tracking_area_code = 1;

    plmn_list = ( { mcc = 208; mnc = 93; mnc_length = 2; } );

    tr_s_preference     = "local_mac"

    ////////// Physical parameters:

    component_carriers = (
      {
      node_function             = "NGFI_RCC_IF4p5";
      node_timing               = "synch_to_ext_device";
      node_synch_ref            = 0;
      frame_type                                          = "FDD";
      tdd_config                                          = 3;
      tdd_config_s                                  = 0;
      prefix_type                                   = "NORMAL";
      eutra_band                                    = 7;
      downlink_frequency                            = 2685000000L;
      uplink_frequency_offset                       = -120000000;
      Nid_cell                                      = 0;
      N_RB_DL                                       = 100;
      Nid_cell_mbsfn                                = 0;
      nb_antenna_ports                              = 1;
      nb_antennas_tx                                = 1;
      nb_antennas_rx                                = 1;
      tx_gain                                          = 90;
      rx_gain                                          = 125;
      pbch_repetition                               = "FALSE";
      prach_root                                    = 0;
      prach_config_index                            = 0;
      prach_high_speed                              = "DISABLE";
      prach_zero_correlation                        = 1;
      prach_freq_offset                             = 2;
```

```
    pucch_delta_shift                                = 1;
    pucch_nRB_CQI                                    = 0;
    pucch_nCS_AN                                     = 0;
    pucch_n1_AN                                      = 0;
    pdsch_referenceSignalPower                              = -27;
    pdsch_p_b                                               = 0;
    pusch_n_SB                                              = 1;
    pusch_enable64QAM                                       = "DISABLE";
    pusch_hoppingMode                                  = "interSubFrame";
    pusch_hoppingOffset                                = 0;
    pusch_groupHoppingEnabled                               = "ENABLE";
    pusch_groupAssignment                                   = 0;
    pusch_sequenceHoppingEnabled                            = "DISABLE";
    pusch_nDMRS1                                       = 1;
    phich_duration                                     = "NORMAL";
    phich_resource                                     = "ONESIXTH";
    srs_enable                                         = "DISABLE";
    /*  srs_BandwidthConfig                                =;
    srs_SubframeConfig                                 =;
    srs_ackNackST                                      =;
    srs_MaxUpPts                                       =;*/

    pusch_p0_Nominal                                   = -96;
    pusch_alpha                                        = "AL1";
    pucch_p0_Nominal                                   = -104;
    msg3_delta_Preamble                                = 6;
    pucch_deltaF_Format1                               = "deltaF2";
    pucch_deltaF_Format1b                              = "deltaF3";
    pucch_deltaF_Format2                               = "deltaF0";
    pucch_deltaF_Format2a                              = "deltaF0";
    pucch_deltaF_Format2b                         = "deltaF0";

    rach_numberOfRA_Preambles                          = 64;
    rach_preamblesGroupAConfig                         = "DISABLE";
    /*
    rach_sizeOfRA_PreamblesGroupA                      = ;
    rach_messageSizeGroupA                             = ;
    rach_messagePowerOffsetGroupB                      = ;
    */
    rach_powerRampingStep                              = 4;
    rach_preambleInitialReceivedTargetPower            = -108;
    rach_preambleTransMax                              = 10;
    rach_raResponseWindowSize                          = 10;
    rach_macContentionResolutionTimer                  = 48;
    rach_maxHARQ_Msg3Tx                                = 4;

    pcch_default_PagingCycle                           = 128;
    pcch_nB                                            = "oneT";
    bcch_modificationPeriodCoeff                            = 2;
    ue_TimersAndConstants_t300                              = 1000;
    ue_TimersAndConstants_t301                              = 1000;
    ue_TimersAndConstants_t310                              = 1000;
    ue_TimersAndConstants_t311                              = 10000;
    ue_TimersAndConstants_n310                              = 20;
    ue_TimersAndConstants_n311                              = 1;
```

```
      ue_TransmissionMode                                    = 1;
      }
    );


    srb1_parameters :
    {
        # timer_poll_retransmit = (ms) [5, 10, 15, 20,... 250, 300, 350, ... 500]
        timer_poll_retransmit    = 80;

        # timer_reordering = (ms) [0,5, ... 100, 110, 120, ... ,200]
        timer_reordering         = 35;

        # timer_reordering = (ms) [0,5, ... 250, 300, 350, ... ,500]
        timer_status_prohibit    = 0;

        # poll_pdu = [4, 8, 16, 32 , 64, 128, 256, infinity(>10000)]
        poll_pdu                 =  4;

        # poll_byte = (kB)
[25,50,75,100,125,250,375,500,750,1000,1250,1500,2000,3000,infinity(>10000)]
        poll_byte                =  99999;

        # max_retx_threshold = [1, 2, 3, 4 , 6, 8, 16, 32]
        max_retx_threshold       =  4;
    }

    # ------- SCTP definitions
    SCTP :
    {
        # Number of streams to use in input/output
        SCTP_INSTREAMS  = 2;
        SCTP_OUTSTREAMS = 2;
    };


    ////////// MME parameters:
    mme_ip_address      = ( { ipv4       = "10.0.0.3";
                              ipv6       = "192:168:30::17";
                              active     = "yes";
                              preference = "ipv4";
                            }
                          );

    enable_measurement_reports = "no";

    ///X2
    enable_x2 = "no";
    t_reloc_prep      = 1000;      /* unit: millisecond */
    tx2_reloc_overall = 2000;      /* unit: millisecond */

    NETWORK_INTERFACES :
    {
        ENB_INTERFACE_NAME_FOR_S1_MME            = "eth0";
        ENB_IPV4_ADDRESS_FOR_S1_MME              = "10.0.0.4/24";
```

```
        ENB_INTERFACE_NAME_FOR_S1U               = "eth0";
        ENB_IPV4_ADDRESS_FOR_S1U                 = "10.0.0.4/24";
        ENB_PORT_FOR_S1U                         = 2152; # Spec 2152

        ENB_IPV4_ADDRESS_FOR_X2C                 = "10.0.0.4/24";
        ENB_PORT_FOR_X2C                         = 36422; # Spec 36422
    };
  }
);

MACRLCs = (
        {
        num_cc = 1;
        tr_s_preference = "local_L1";
        tr_n_preference = "local_RRC";
        }
);

L1s = (
        {
        num_cc = 1;
        tr_n_preference = "local_mac";
        }
);

RUs = (
    {
        local_if_name  = "eth1";
        remote_address = "10.10.10.111";
        local_address  = "10.10.10.153";
        local_portc    = 50002;
        remote_portc   = 50002;
        local_portd    = 50003;
        remote_portd   = 50003;
        local_rf       = "no"
        tr_preference  = "udp_if4p5"
        nb_tx          = 1
        nb_rx          = 1
        att_tx         = 0
        att_rx         = 0;
        eNB_instances  = [0];
        is_slave       = "no"
    }
);

THREAD_STRUCT = (
  {
    #three config for level of parallelism "PARALLEL_SINGLE_THREAD",
"PARALLEL_RU_L1_SPLIT", or "PARALLEL_RU_L1_TRX_SPLIT"
    parallel_config    = "PARALLEL_RU_L1_TRX_SPLIT";
    #two option for worker "WORKER_DISABLE" or "WORKER_ENABLE"
    worker_config      = "WORKER_ENABLE";
  }
);
```

```
log_config = {
      global_log_level                    ="info";
      global_log_verbosity                ="medium";
      hw_log_level                        ="info";
      hw_log_verbosity                    ="medium";
      phy_log_level                       ="info";
      phy_log_verbosity                   ="medium";
      mac_log_level                       ="info";
      mac_log_verbosity                   ="high";
      rlc_log_level                       ="info";
      rlc_log_verbosity                   ="medium";
      pdcp_log_level                      ="info";
      pdcp_log_verbosity                  ="medium";
      rrc_log_level                       ="info";
      rrc_log_verbosity                   ="medium";
};
```

E) Established VPN tunnel in the same VLAN-100 segment as the SR-IOV networks. The server running at the same physical host as the 5G Core Network. The tun0 interface directly maps the routes from the eno2 interface, which is used by the SR-IOV driver to run the Core Network container.

```
eno1: flags=4419<UP,BROADCAST,RUNNING,PROMISC,MULTICAST>  mtu 9000
        inet 128.39.120.239  netmask 255.255.254.0  broadcast 128.39.121.255
        inet6 fe80::d294:66ff:fe09:4f2f  prefixlen 64  scopeid 0x20<link>
        inet6 2001:700:700:22:d294:66ff:fe09:4f2f  prefixlen 64  scopeid 0x0<global>
        ether d0:94:66:09:4f:2f  txqueuelen 1000  (Ethernet)
        RX packets 93900324  bytes 49089459229 (49.0 GB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 89889878  bytes 50061554230 (50.0 GB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 68  memory 0x95000000-957fffff

eno2: flags=4419<UP,BROADCAST,RUNNING,PROMISC,MULTICAST>  mtu 9000
        inet 10.0.0.1  netmask 255.255.255.0  broadcast 10.0.0.255
        inet6 fe80::d294:66ff:fe09:4f31  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:09:4f:31  txqueuelen 1000  (Ethernet)
        RX packets 220297826  bytes 1923349453206 (1.9 TB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 95919924  bytes 82406524244 (82.4 GB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x94000000-947fffff

eno3: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        ether d0:94:66:09:4f:33  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x93000000-937fffff

eno4: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 9000
        inet 192.168.1.1  netmask 255.255.255.0  broadcast 192.168.1.255
        inet6 fe80::d294:66ff:fe09:4f35  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:09:4f:35  txqueuelen 1000  (Ethernet)
        RX packets 1137466  bytes 97506897 (97.5 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 74235  bytes 5843081 (5.8 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 103  memory 0x92000000-927fffff

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 26  bytes 2688 (2.6 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 26  bytes 2688 (2.6 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

tun0: flags=4305<UP,POINTOPOINT,RUNNING,NOARP,MULTICAST>  mtu 9000
        inet 10.8.0.1  netmask 255.255.255.0  destination 10.8.0.1
        inet6 fe80::10e5:342e:7e0f:e730  prefixlen 64  scopeid 0x20<link>
        unspec 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00  txqueuelen 100  (UNSPEC)
        RX packets 78  bytes 9733 (9.7 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 104  bytes 10645 (10.6 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

root@atlas:~#
```

F) A VPN tun0 interface at the Centralized Unit node, connecting to the 10.8.0.1 endpoint, via the same SR-IOV VLAN-100 network, through the 10.0.0.2 interface, which is a direct fiber link to the Core Network host. The tun0 maps routes from the eno1 interface, which is used by the Centralized Unit as a entrypoint for the Docker communication via the SR-IOV passthrough.

```
root@poaseidon:~# ifconfig
docker0: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        inet 172.17.0.1  netmask 255.255.0.0  broadcast 172.17.255.255
        ether 02:42:82:ad:65:19  txqueuelen 0  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

eno1: flags=4419<UP,BROADCAST,RUNNING,PROMISC,MULTICAST>  mtu 9000
        inet 10.0.0.2  netmask 255.255.255.0  broadcast 10.0.0.255
        inet6 fe80::d294:66ff:fe02:17b1  prefixlen 64  scopeid 0x20<link>
        ether d0:94:66:02:17:b1  txqueuelen 1000  (Ethernet)
        RX packets 95929085  bytes 82407533913 (82.4 GB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 220291644  bytes 1923348719191 (1.9 TB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 68  memory 0x95000000-957fffff

eno2: flags=4099<UP,BROADCAST,MULTICAST>  mtu 9000
        ether d0:94:66:02:17:b3  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x94000000-947fffff

eno3: flags=4099<UP,BROADCAST,MULTICAST>  mtu 1500
        ether d0:94:66:02:17:b5  txqueuelen 1000  (Ethernet)
        RX packets 0  bytes 0 (0.0 B)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 0  bytes 0 (0.0 B)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
        device interrupt 82  memory 0x93000000-937fffff

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 204357  bytes 17193906 (17.1 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 204357  bytes 17193906 (17.1 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

tun0: flags=4305<UP,POINTOPOINT,RUNNING,NOARP,MULTICAST>  mtu 9000
        inet 10.8.0.2  netmask 255.255.255.0  destination 10.8.0.2
        inet6 fe80::bd7b:d5eb:6560:f703  prefixlen 64  scopeid 0x20<link>
        unspec 00-00-00-00-00-00-00-00-00-00-00-00-00-00-00-00  txqueuelen 100  (UNSPEC)
        RX packets 78  bytes 8765 (8.7 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 80  bytes 9593 (9.5 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```