UNIVERSITY
OF OULU

FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING
DEGREE PROGRAMME IN ELECTRONICS AND COMMUNICATIONS ENGINEERING

MASTER'S THESIS

5G NETWORK END-TO-END DELAY MEASUREMENTS FOR LIVE VIDEO STREAMING

Author	Antti Pauanne
Supervisor	Ari Pouttu
Second Examiner	Ville Niemelä

February 2020

Pauanne A. (2020) 5G Network End-to-End Delay Measurements for Live Video Streaming. University of Oulu, Faculty of Information Technology and Electrical Engineering, Degree Programme in Electronics and Communications Engineering. Master's Thesis, 73 p.

ABSTRACT

Focus of this thesis is in the data transmission delay comparison between Edge server and Cloud server when utilizing either 4G or 5G connectivity. In previous mobile phone network generations for example a multimedia server had to be installed on a Cloud server in the internet. 5G mobile phone network introduces a new concept called Edge server. Edge server is located close to the base station and therefore it is assumed to shorten the data transmission delay between the 5G mobile/client and a server application. Edge server can be used both in 4G and 5G networks.

In this thesis first the 5G network and the essential new 5G architecture main design principles are gone through. Next the 5G Test Network that is used as a test environment is described and 5G main modules like Multi-access Edge Computing are introduced. 5G performance is clarified and compared against 4G.

Delay testing is done in the 5G Test Network using Hospital Use Case demo. There operating room personnel like doctors and nurses is wearing Augmented Reality glasses and they are streaming their view together with patient status related information to multimedia server residing in 5G Test Network Edge server or in internet cloud. From the multimedia server the video is streamed by for example students, medical experts or consultants in a remote location. As part of the thesis the test system is defined and built based on the Hospital Use Case demo. Test specification is created, and tests are executed according to it. Results are recorded and analysed.

Data transmission delays between the video stream originator and multimedia server are measured using Qosium measurement system. Also delay between the multimedia server and the streaming client is measured. Measurements are done for configurations where multimedia server is located at the Edge server and the internet cloud server. Both 4G and 5G connectivity is used for both server locations.

When delay measurement results were compared it became clear that Edge server has much shorter data transmission delays compared to the internet cloud server. With 5G connectivity the delay was measured to be around 10 milliseconds for both uplink and downlink. With internet cloud the delays varied between 31 and 45 milliseconds with 5G connection. It can be concluded that from today's mobile phone networks, 5G network does offer the fastest connection to a server environment by utilizing Edge server.

Key words: MEC, Edge server, latency.

Pauanne A. (2020) 5G verkon viiveen mittaaminen videostriimille. Oulun yliopisto, tieto- ja sähkötekniikan tiedekunta, elektroniikan ja tietoliikennetekniikan tutkinto-ohjelma. Diplomityö, 73 p.

TIIVISTELMÄ

Tämä diplomityö keskittyy vertaamaan datatiedonsiirron eroja reunapalvelimen ja internetin pilvipalvelimen välillä 4G ja 5G matkapuhelinverkossa. Aiempien sukupolvien matkapuhelinverkoissa esimerkiksi multimediapalvelin oli asennettava internetin pilvipalvelimelle. Viidennen sukupolven matkapuhelinverkossa otetaan käyttöön reunapalvelin. Reunapalvelin sijaitsee tukiaseman läheisyydessä ja täten sen oletetaan lyhentävän 5G-päätelaitteen ja palvelimen sovelluksen välistä tiedonsiirtoviivettä. Reunapalvelinta voidaan käyttää sekä neljännen että viidennen sukupolven matkapuhelinverkoissa.

Tässä diplomityössä käydään ensin läpi 5G-matkapuhelinverkko ja sen arkkitehtuurin pääsuunnittelukriteerit. Seuraavaksi kuvataan testaamisessa käytettävä 5G-testiverkko ja 5G-verkon tärkeimmät moduulit kuten Multi-access Edge Computing. 5G-verkon suorituskyky selitetään ja sitä verrataan edelliseen 4. sukupolven verkkoon.

Viivemittaukset tehdään 5G testiverkossa käyttäen 5G lääketieteen käyttötapauksen demoympäristöä. Siinä operointihuoneen henkilöstöllä, kuten lääkäreillä ja hoitajilla, on yllään lisätyn todellisuuden lasit. Lasit lähettävät henkilön näkymän ja potilaaseen liittyvää tietoa 5G-testiverkon reunapalvelimella tai internetin pilvipalvelimella sijaitsevalle multimediapalvelimelle. Multimediapalvelimelta video striimataan esimerkiksi lääketieteen opiskelijoille, asiantuntijoille tai konsulteille, jotka ovat etäällä lähetyspaikasta. Osana diplomityötä määritellään ja rakennetaan lääketieteen käyttötapauksen demon perustuva testausjärjestelmä. Testispesifikaatio luodaan, testit suoritetaan sen perusteella. Testitulokset tallennetaan ja analysoidaan.

Tiedonsiirtoviiveet videolähteen ja multimediapalvelimen välillä mitataan käyttäen Qosium mittausjärjestelmää. Myös multimediapalvelimen ja videostriimin vastaanottajan väliset viiveet mitataan. Mittaukset tehdään konfiguraatiolle, jossa multimediapalvelin on sijoitettu reunapalvelimelle ja konfiguraatiolle, jossa se on sijoitettu internetin pilvipalvelimelle. Sekä 4G että 5G-yhteyttä käytetään molemmille konfiguraatiolle.

Kun mittaustuloksia verrataan, käy selväksi, että reunapalvelimella on huomattavasti lyhyempi tiedonsiirtoviive kuin internetin pilvipalvelimella. 5G-yhteydellä mitattu viive oli noin 10 ms sekä ylössyöttö- että alassyöttösuuntaan. Internetin pilvipalvelimella viiveet vaihtelivat 31 ja 45 millisekunnin välillä 5G-yhteydellä. Voidaankin todeta, että nykyisistä matkapuhelinverkoista 5G-verkko tarjoaa nopeimman yhteyden palvelinympäristöön reunapalvelimen avulla.

Avainsanat: MEC, reunapalvelin, viive.

TABLE OF CONTENTS

FOREWORD

This diploma thesis has been done in the Centre for Wireless communications (CWC), University of Oulu, from summer 2019 to February 2020. It has been extremely interesting to finalize my studies that were on hold for over 20 years as I worked in the telecommunications industry. I had the chance to again pursue the Master of Science (MSc) degree that was so close but still so far. Returning and getting this thesis approved is the final step before the goal is achieved.

First of all, I want to thank my supervisor professor Ari Pouttu for giving me this opportunity to do my thesis at CWC by hiring me and believing in me. Postdoctoral researcher Ville Niemelä who is the second examiner of this thesis has been my everyday support and he has done a huge work by guiding me and by reviewing the thesis as it has been progressing. His help has been invaluable. Olli Liinamaa, who works as 5G Test Network (5GTN) project manager is to thank for offering me the possibility to do this thesis as part of his project and for encouraging me along the way.

I also need to thank personnel of CWC for the support I have received. There has been a lot of persons helping me in details of 5G Test Network, writing the thesis and project work. These persons include Jari Moilanen, Vesa Halonen, Kenichi Komatsu and Rainer Laaksonen. Thanks also for the rest of the CWC personnel for creating a work environment and atmosphere that is both welcoming and inspiring.

For more than anything I need to thank my family for the endless support the have given me. They have been the corner stone of my life and encouraged me on my long journey towards the MSc degree. My wife Raila has never stopped believing in me and has gently pushed me forward to reach my dreams.

Ii, 3 February, 2020

Antti Pauanne

LIST OF ABBREVIATIONS AND SYMBOLS

2G	Second Generation Mobile Phone Network
3G	Third Generation Mobile Phone Network
3GPP	Third Generation Partnership Project
3.5G	A version of the 3G network
4G	Fourth Generation Mobile Phone Network
5G	Fifth Generation Mobile Phone Network
6G	Sixth Generation Mobile Phone Network
5G PPP	5G Infrastructure Public Private Partnership
5GC	5G Core
5GTN	5G Test Network
AF	Application Function
AM	Ante Meridiem
AMF	Access and Mobility Management Function
API	Application Programming Interface
APN	Access Point Name
AR	Augmented Reality
ARIB	Association of Radio Industries and Businesses
ATIS	Alliance for Telecommunications Industry Solutions
AUSF	Authentication Server Function
BTS	Base Transceiver Station
CAPEX	Capital Expenditure
CCSA	China Communications Standards Association
CDN	Content Delivery Network
CGF	Charging Gateway Function
CRAN	Cloud RAN
CUPS	Control/User Plane Separation
CWC	Centre for Wireless Communications
D2D	Device to Device
DHCP	Dynamic Host Configuration Protocol
DPI	Deep Packet Inspection
E2E	End-to-End
EB	Exabyte
eMBB	enhanced Mobile Broadband
eMTC	enhanced Machine Type Communications
eNB	evolved Node B
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
eV2X	enhanced Vehicle-to-Everything
fps	frames per second
Gb	Gigabit
Gbps	Gigabit per second
GHz	Gigahertz
GPRS	General Packet Radio Service
GTP	GPRS Tunnelling Protocol
GTP-U	GTP-User plane

HLS	HTTP Live Stream
HRLLC	Highly Reliable Low Latency Communication
HSDPA	High Speed Downlink Packet Access
HSS	Home Subscriber Server
HTTP	Hypertext Transfer Protocol
HW	Hardware
Hz	Hertz
IIoT	Industrial IoT
IoT	Internet of Things
IMS	IP Multimedia System
IMT	International Mobile Telecommunications
IP	Internet Protocol
IPTV	Internet Protocol Television
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union Radiocommunication Sector
Km2	Square Kilometre
Km/h	Kilometres per hour
kbit/s	kilobit per second
KPI	Key Performance Indicator
LAN	Local Area Network
LIG	Lawful Intercept Group
LTE	Long Term Evolution
LTS	Long Term support
M2M	Machine to Machine
Mbit	Megabit
Mbps	Megabit per second
MC	Mission Critical
MCPTT	Mission Critical Push to Talk
MIMO	Multiple-Input and Multiple-Output
MEC	Multi-access Edge Computing
MME	Mobility Management Entity
mMTC	massive Machine-Type Communications
mmWave	millimetre Wave
ms	millisecond
MSc	Master of Science
MTC	Machine-Type of Communications
NAT	Network Address Translation
NAS	Network-Attached Storage
NB-IoT	Narrow Band IoT
NEF	Network Exposure Function
NF	Network Function
NFV	Network Function Virtualisation
NFVI	NFV Infrastructure
NR	New Radio
NRF	NF Repository Function
NR-U	NR Unlicensed spectrum
NSA	Non-Stand Alone
OBS	Open Broadcaster Software

OPEX	Operational Expenditure
OS	Operating system
PC	Personal Computer
PCF	Policy Control Function
PCRF	Policy and Charging Rules Function
PDN	Packet Data Network
PGW	PDN Gateway
PM	Post Meridiem
PTP	Precision Time Protocol
PTPd	Precision Time Protocol daemon
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RNIS	Radio Network Information Service
RTMP	Realtime Messaging Protocol
s	second
SA	Stand-Alone
SBA	Service Based Architecture
SDN	Software Defined Networking
SGi	name for PGW interface
SGW	Serving Gateway
SGW-LBO	SGW with Local Breakout
SIM	Subscriber Identity Module)
SMF	Session Management Function
SW	Software
TSDSI	Telecommunications Standards Development Society in India
TTA	Telecommunications Technology Association
TTC	Telecommunication Technology Committee
TV	Television
UDN	Ultra Dense Network
UDM	Unified Data Management
UE	User Equipment
UPF	User Plane Function
URLLC	Ultra Reliable Low Latency Communication
USB	Universal Serial Bus
V2X	Vehicle-to-Everything
VLC	VideoLAN Client
VM	Virtual Machine
vMEC	Virtual MEC
VNF	Virtual Network Function
VPN	Virtual Private Network
VR	Virtual Reality
WiFi	Wireless Fidelity
WLAN	Wireless Local Area Network

1 INTRODUCTION

Digital mobile networks have been evolving with a cycle of approximately 10 years. In 1990’s there was the 2nd generation mobile phone network (2G), first decade of 21st century 3rd generation mobile phone network (3G) was introduced and early 2010s it was the time for the 4th generation (4G) of the mobile phone network and Long Term Evolution (LTE). At the time of writing this thesis in the late 2019 and the early 2020, previous generations still exist and some also evolve but 5G networks are now being defined and built. [1]

Each of the mobile network generations must answer to certain needs that the previous generation has not been able to fulfil. As an example, the 3G was to enable mobile internet for the consumer devices. All of the needs may not be fulfilled in the first version of each network, but they do continue to develop to better meet the needs that are either evolving or emerging during the lifecycle of each of the mobile network generations. [2]

Table 1 presents main services, differentiators and weaknesses for all mobile phone network generations up to 4G [2].

Table 1. Evolution of technology generations in terms of services and performance

Generation	Primary services	Key differentiator (against previous generations)	Weakness (addressed by subsequent generation)
1G	Analogue phone calls	Mobility	Poor spectral efficiency, major security issues
2G	Digital phone calls and messaging	Secure, mass adoption	Limited data rates – difficult to support demand for internet/e-mail
3G	Phone calls, messaging, data	Better internet experience	Real performance failed to match hype, failure of WAP for internet access
3.5G	Phone calls, messaging, broadband data	Broadband internet, applications	Tied to legacy, mobile specific architecture and protocols
4G	All-IP services (including voice, messaging)	Faster broadband internet, lower latency	Millisecond level latency, up to 1Gb/s data rate

In Table 1 WAP stands for Wireless Application Protocol and IP for Internet Protocol.

3.5G is a version of 3G network introducing High Speed Downlink Packet Access (HSDPA) and rising the data speeds up to 10x faster than in the original 3G release. 3.5G together with 4G has brought faster data speeds and lower latencies available to consumers leading to a change in the way customers use their devices. [2] It has been estimated that total data traffic generated by smartphones is to grow 45% annually, leading to smartphone traffic to increase 10 times and total mobile data traffic for all device to grow by eight times between 2016 and 2022 [3]. When the growth is that fast, the capacity offered by the 3.5G and 4G will be used up quickly. A new solution is needed and 5G is to bring more bandwidth among other features. The extremely fast growth in mobile phone data is visible in Figure 1 below [4].

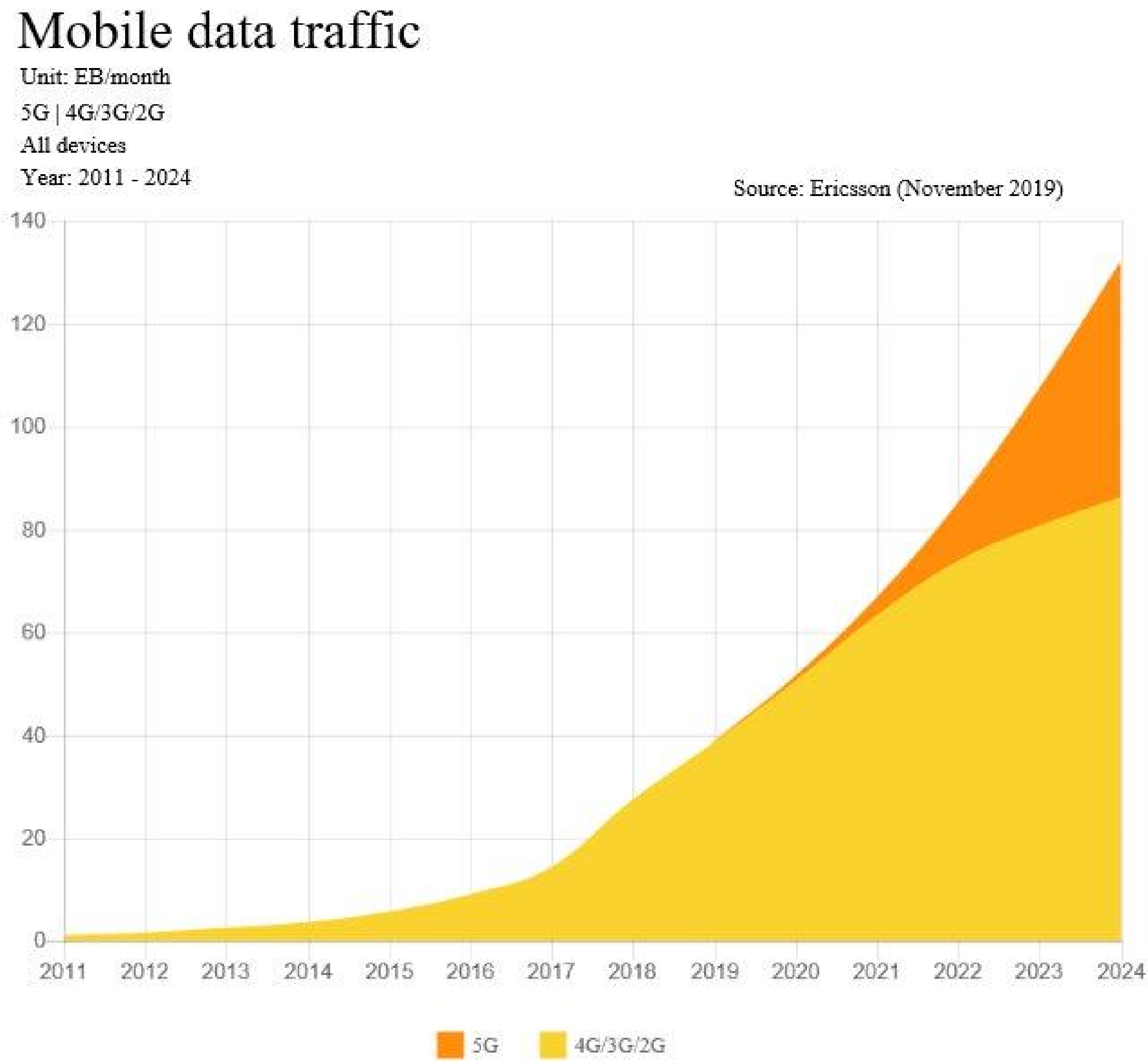


Figure 1. Global mobile data traffic 2011 – 2024 in exabytes (EB) per month. Based on the data available 7-Jan-2020.

The number of users in mobile networks is growing. The growth does not come from the regular consumers' mobile phone usage but mainly from Internet of Things (IoT) as more and more devices will have intelligence and are connected to the internet. Therefore, Machine to machine (M2M) and Device to Device (D2D) communication will be growing fast. Ericsson has forecasted that the amount of IoT devices will pass the amount of all other connected devices already during 2019. By 2024 there will be 34 billion connected devices and 66% of those will be IoT devices. Trend is visible in Figure 2. [4]. IoT is one of the main focuses of 5G that is promising the capacity, reliability and ultra-low latency required for mission critical services and also for the growth of massive IoT [5].

IoT- as name says- is about connecting all possible devises to the internet. These devices can be for example home appliances, cars, media devices, sensors, etc. These devices can collect information and send it through internet like sensors do or they can receive information through internet and then act upon it like for example cars could do, or they can do both, send and receive information as electronic heating systems do.

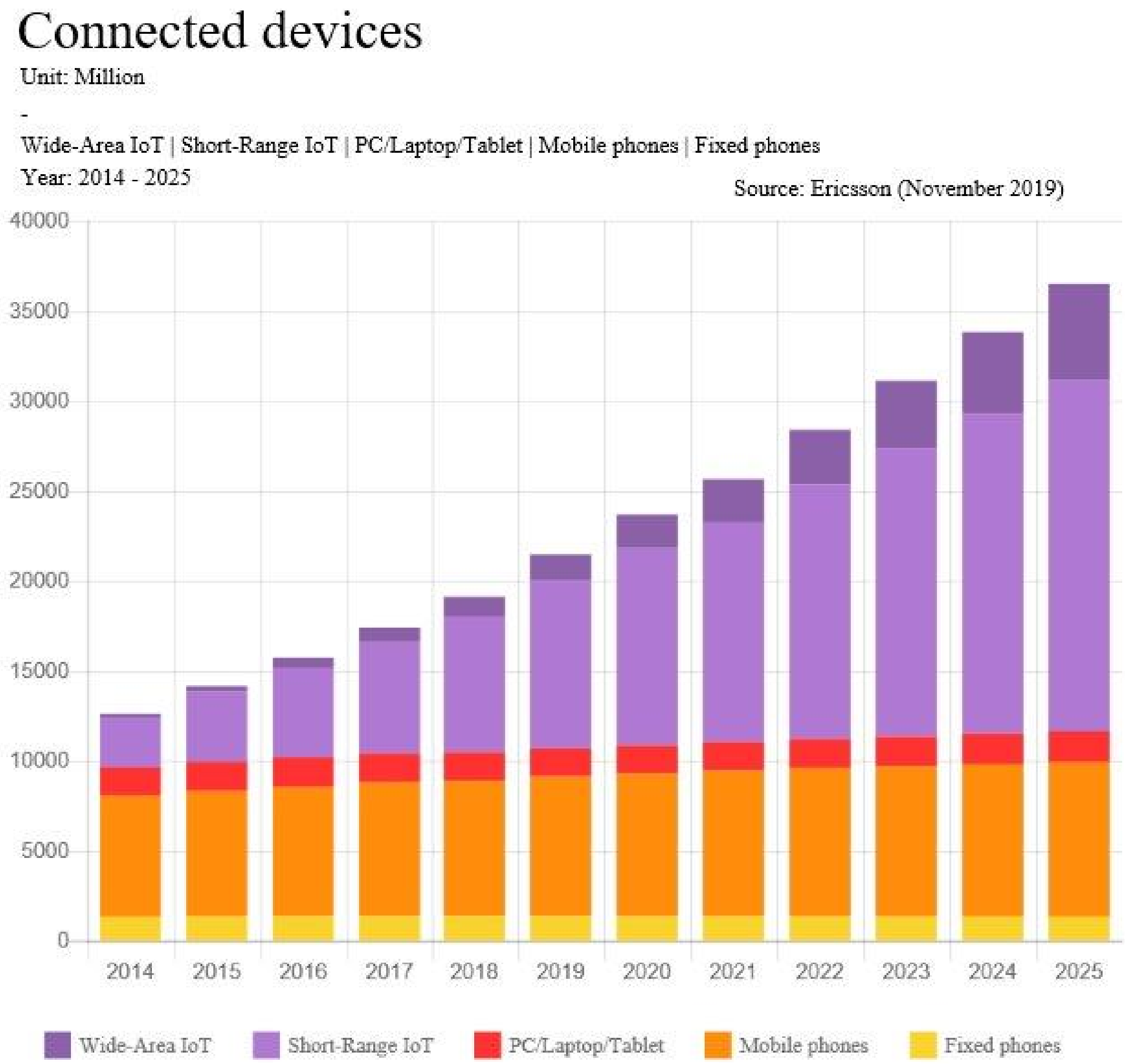


Figure 2. Global connected devices from 2014 to 2024. Based on the data available 7-Jan-2020.

In Figure 2 PC stands for Personal Computer.

As seen from Figures 1 and 2 both number of connected devices and the amount of mobile data traffic is increasing very fast. Current mobile phone network generations will be in difficulties to handle the increasing number of devices and traffic. From 2019 onwards new mobile phone network generation 5G, that has been under development for several years, is introduced. It will extend the data capacity and the user experienced throughput of the mobile phone networks, connect extremely high number of equipment and increase the reliability of the connections by introducing low latency ultra-reliable connections. [1]

One of the main promises of the new 5G network is low latency. That low latency is utilized in a new Edge server technology introduced as one major new technology of the 5G network. It adds computing power, that was previously available only at internet cloud, close to the base station. This gives service providers opportunity to open new vertical markets by supporting new services and applications. One of the main features of the Edge server technology is the low latency communication between the client and the Edge server. Focus of this thesis is to understand the possible benefits the Edge server technology has in the data transmission delay area. Edge server and Cloud server compared against each other by measuring communication delays between the mobile network client and both servers. Both 4G and 5G connectivity will be used. [6]

2 5G NETWORK

2.1 5G specification releases

5G specifications are defined by the 3rd Generation Partnership Project (3GPP) which is pulling together several telecommunications standard development organizations (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC) that are contributing and participating in specification creation. ARIB stands for The Association of Radio Industries and Businesses located in Japan. ATIS is abbreviation for The Alliance for Telecommunications Industry Solutions in the United States of America. CCSA means China Communications Standards Association and ETSI is The European Telecommunications Standards Institute that has its main site at Sophia Antipolis, France. TSDSI means Telecommunications Standards Development Society in India. TTA is abbreviation for Telecommunications Technology Association in Korea and TTC for Telecommunication Technology Committee from Japan. [7]

3GPP is using a system of parallel releases when defining 5G specifications. By having parallel releases developers have access to a stable release at the same time as new functionalities are being defined for next releases. 3GPP releases from 15 to 17 has been defined to be 5G releases. Release 15 status is frozen as functional freeze happened March 22nd, 2019 and the end date when protocol was stable was the 7th of June 2019. Release 16 is now under development and release 17 is now at early definition state that leads to Approval of Rel-17 content as visible in Figure 3. [8] Roadmap for 3GPP releases is visible in Figure 3 [9].

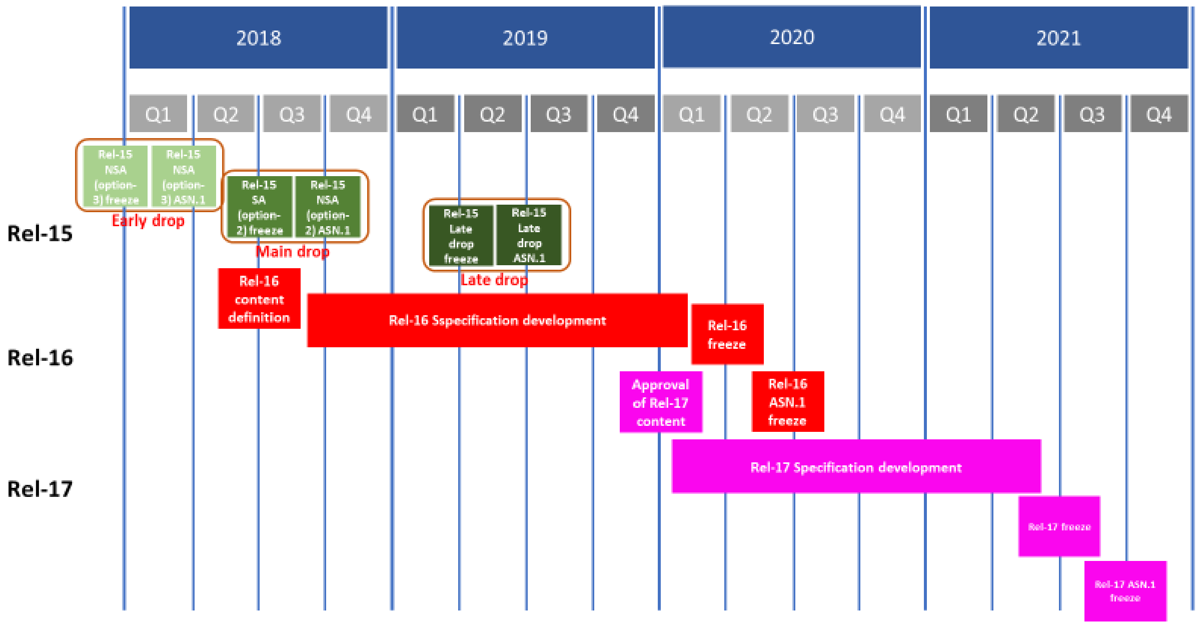


Figure 3. 3GPP Release roadmap.

Release 15 is mainly focusing on the initial phase of the 5G. 5G is defined in different phases: phase 1 is defined in Release 15 and following releases will contain specifications for the coming phases. Other Rel-15 features include further enhancements on Critical Communications (including Ultra Reliable Low Latency Communication (URLLC) and

Highly Reliable Low Latency Communication (HRLLC)), Machine-Type of Communications (MTC) and IoT, Vehicle-to-Everything (V2X), Mission Critical (MC), and features related to Wireless Local Area Network (WLAN) and unlicensed spectrum. Rel-15 also introduces Non-Stand Alone (NSA) and Stand-Alone (SA) architectures that are important for this diploma thesis as 5G Test Network [10] that is used in this work uses NSA. [11]

Release 16 specification development work is on-going and specification freeze is planned to happen during the first half of 2020. Main features of Rel-16 include advanced use cases beyond LTE V2X, industrial IoT and URLLC enhancements, 5G New Radio (NR) operation in unlicensed bands. Also, several system improvements and enhancements in the areas of positioning, Multiple-Input and Multiple-Output (MIMO) and Power consumption are under specification.[9]

Release 17 is in early phase of planning but the work areas that are known include NR light, Small data transfer optimization, Sidelink enhancement, NR above 52.6 Gigahertz (GHz), Multi Subscriber Identity Module (Multi SIM) operation, NR multicast broadcast. Enhancements are planned for the areas in coverage, Narrow Band IoT (NB-IoT) and enhanced Machine Type Communications (eMTC), Industrial IoT (IIoT), MIMO, NR for Non-Terrestrial Networks, Integrated Access and Backhaul, NR Unlicensed spectrum (NR-U), Power saving, Radio Access Network (RAN) data collection and positioning. [9]

2.2 5G main use cases and requirements

There are several use cases defined for 5G. A few main use cases, relevant for this thesis, are presented here. Use cases can be classified, for example, into four categories [12] [13]:

- Enhanced Mobile Broadband (eMBB)
  - Enabled by higher data rates together with lower latency
  - Lower cost per data unit
  - 1 – 20 Gigabits per second (Gb/s)
  - Needed for example by Augmented Reality (AR) and Virtual Reality (VR)
- Massive Machine-Type Communications (mMTC)
  - Tens of billions of devices expected to be connected to internet by 2020
  - Needed by for example Industrial IoT
- Enhanced Vehicle-to-Everything (eV2X)
  - Enabled by low latency
  - Extremely high reliability requirements
  - Vehicles communicating to each other and other devices around them
  - Autonomous Driving, safety enhancements
- Ultra-Reliable and Low Latency Communications URLLC
  - Need for very reliable and low latency network
  - New services like remote control, automation, robotics and drone control enabled

In this thesis focus is mainly on eMBB as the used set-up includes real time AR, VR and video streaming.

Use cases together with architectural key design principles and 5G vision leads to requirements that 5G system needs to fulfil. List of key performance requirements for the 5G system below. [14]

- 10000 times more traffic

- 10 – 100 times more devices
- Below 1 millisecond (ms) latency
- 10 years M2M battery life
- M2M ultra low cost
- Reduced power consumption and increased battery life
- Over 10 Gbit/s peak data rates
- 100 megabits per second (Mbit/s) whenever needed

From the above requirements, this diploma thesis focuses especially on the latency requirement of less than one millisecond.

2.3 5G Architecture

5G architecture is based on the requirement of recursive structure. There is a requirement for slices where end-to-end 5G network functionality can be split into several separate independent instances. Each of these instances offer 5G service for certain customer or functionality. This leads to a requirement of recursive structure. In the recursive structure the architecture can support several instantiated services by building new services out of the same services. In other words, same 5G network functionality can support several slices at the same time. 5G recursive model architecture is presented in Figure 4. [15]

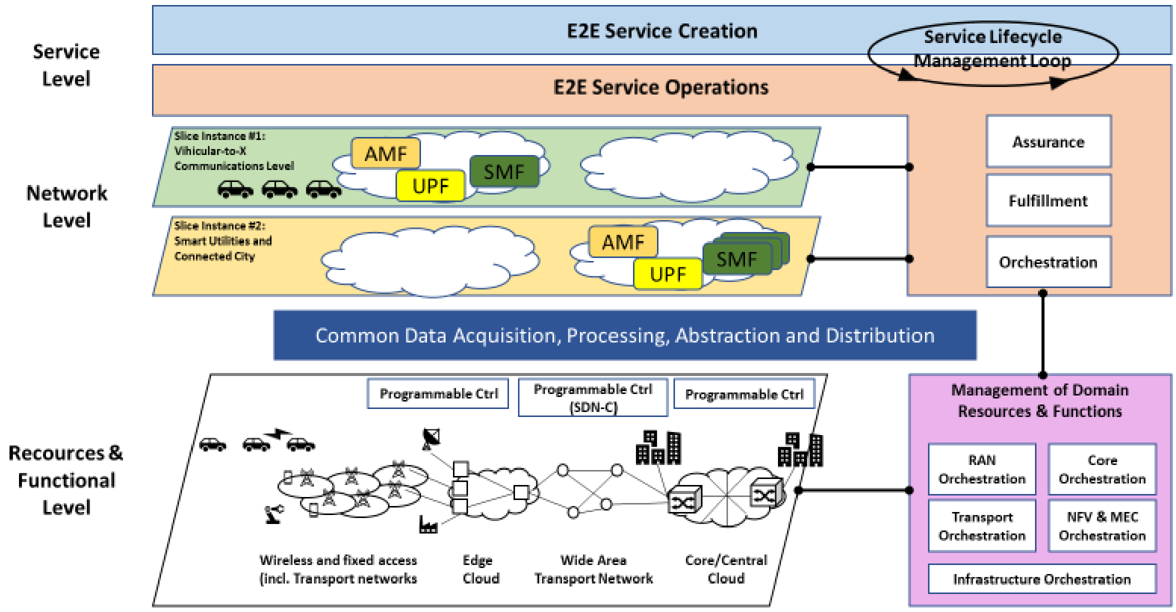


Figure 4. 5G Overall Recursive Model Architecture.

All the phases of Service Level lifecycle management in Service Creation and Service Operations are done by closed-loop functions of Assurance, Fulfilment and Orchestration. On the Network Level E2E (End-to-End) Service Operations functions communicate with different Network Slices. E2E Service Operations is also responsible for the communication with Management of Domain Resources & Functions modules. These modules include RAN Orchestration, Core Orchestration, Transport Orchestration, NFV (Network Function

Virtualisation) & MEC (Multi-access Edge Computing) Orchestration and Infrastructure Orchestration. These functions are then on the Resources and Functional Level communicating with Radio Networks, Edge Cloud, Wide Area Transport Network and Core Central Cloud. There are also common platform functions that can be accessed from all the levels of the 5G system: Common Data Acquisition, Processing, Abstraction and Distribution functions. [15]

2.4 5G architecture main design principles

2.4.1 Slices

Network slicing is one of the main design concepts of the 5G. Purpose of the network slicing is to create a way for the operators to offer their customers a part of their network for a specific use case. This way each of the use cases can have a specific subset of 5G network resources tailored for their needs. Network slicing is one form of virtual networking: slices are virtual 5G networks built from same 5G network modules supporting several slices (virtual networks) efficiently at the same time. An example of 5G network slicing is presented below in Figure 5. [16]

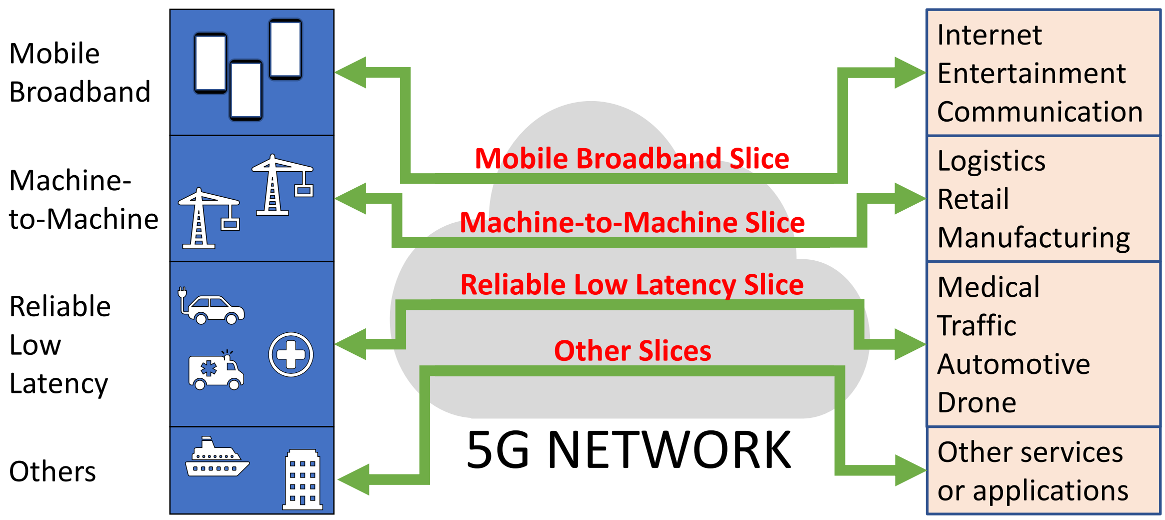


Figure 5. 5G network slicing example.

Software Defined Networking (SDN) and Network Function Virtualization (NFV) together with cloud technologies are the key elements enabling slicing. With these technologies network is softwarizated and key properties like programmability, flexibility and modularity that are needed for network slicing are provided. Virtualization technologies are needed that both Hardware (HW) and Software (SW) resources can be offered to services needed by the slices. [17] [18]

By using slicing, operators can efficiently utilize their network by supporting multiple virtual networks in their common 5G network. Operator can efficiently offer only the functionality the customer, service or use case needs without unnecessarily utilizing resources that are not needed for the particular case. This way operators are capable of efficiently utilize their network and at the same time enable service differentiation, which is one of the benefits of 5G. [18]

2.4.2 Software Defined Networking

Software defined networking is one of the most important technology enablers of 5G. The main idea of SDN is to separate control plane of the network from the data plane and to offer centralized control of the network. [19]

There are four main innovations in SDN [20]:

1. Separation of the control and data planes
2. Centralization of the control plane
3. Programmability of the control plane
4. Standardization of Application Programming Interfaces (APIs)

Simplified structure of the SDN architecture is presented in Figure 6 below [20].

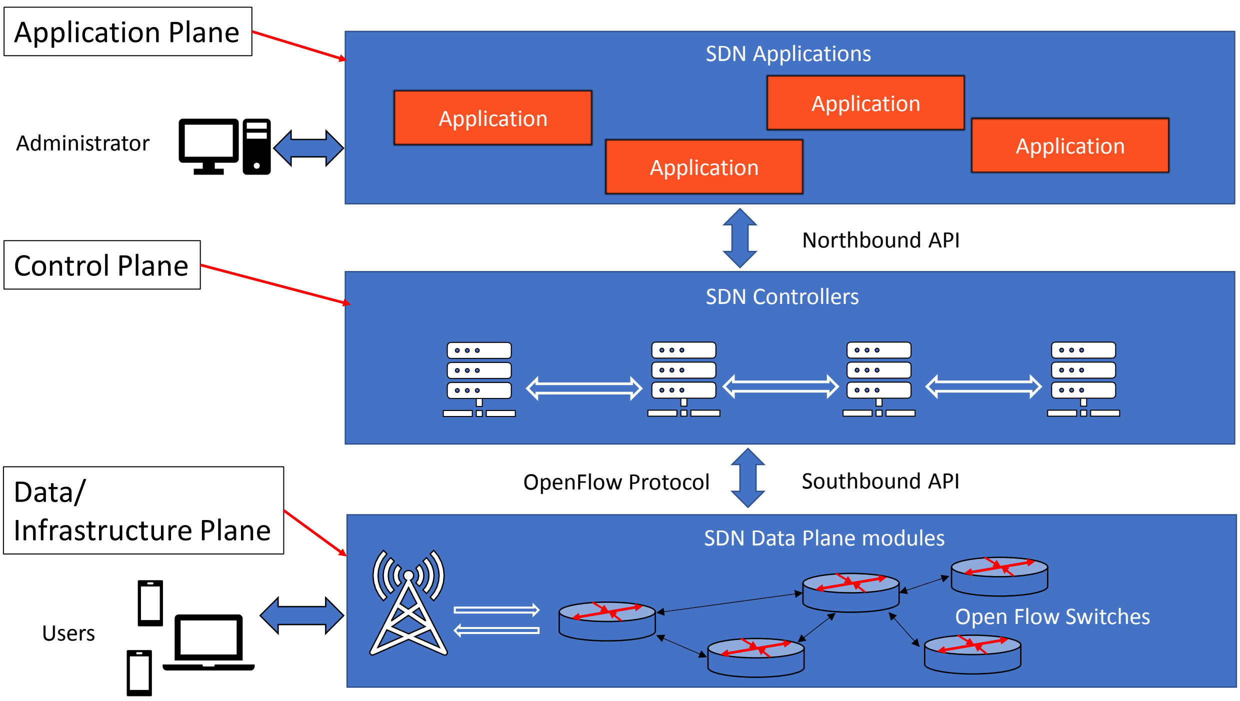


Figure 6. SDN architecture.

All user data is transported at the data plane. Network needs to find best possible way to forward the data to its destination. After deciding the best way to route the data packets, control plane needs to prepare forwarding tables and send them to the data plane forwarding elements like Open Flow switches. Switches then route the data packets as instructed by the forwarding tables, they receive from the control plane controllers. This way the control logic creation is separated from the data plane and switches are just implementing the routing instructions they receive from controller greatly simplifying the switch complexity and reducing the cost of each switch. [20]

Controllers send the control messages to data plane switches through Southbound API. Southbound API defines the communication message structure between control and data planes. Southbound API is Open Flow interface standardized by the Open Network Foundation. [20]

Control plane also keeps track of different network statistics like traffic related items and network HW status. This information is needed by the network manager that is administrating the network and does the changes to network configuration and policies. [20]

Northbound API is the interface used to communicate between SDN Application Plane and Control Plane. These API's are used to send communication from applications to SDN controllers and vice versa. They are also used to integrate SDN Controllers with network automation stacks and network orchestration. [21] Northbound APIs have not been standardized yet but there are several open source projects and groups dedicated to developing northbound API. Several different API definitions have been created for different application needs and each controller can have a different programming interface [20][21].

SDN is not a new technology developed for 5G, but it is a generic networking technology that can be used also in 5G. It is mainly based on open software principle and therefore the definition of it is not done as part of the 5G specification operations but as part of open software definition.

2.4.3 Network Function Virtualization

Network function virtualization improves the flexibility of network service provisioning and shortens the time that it takes for the new services to be deployed to market. NFV virtualizes previously HW based solutions with SW implementation using virtualization techniques and running the SW on general purpose off-the-self programmable HW like x86 based servers. Several Virtual Network Functions (VNFs) can share the same HW resource and they can be run at the same time. This way NFV lowers the operator's capital expenditure (CAPEX) and operational expenditure (OPEX). Figure 7 presents an example of both conventional dedicated HW based components approach and NFV based approach of a communications network. [22][23][24]

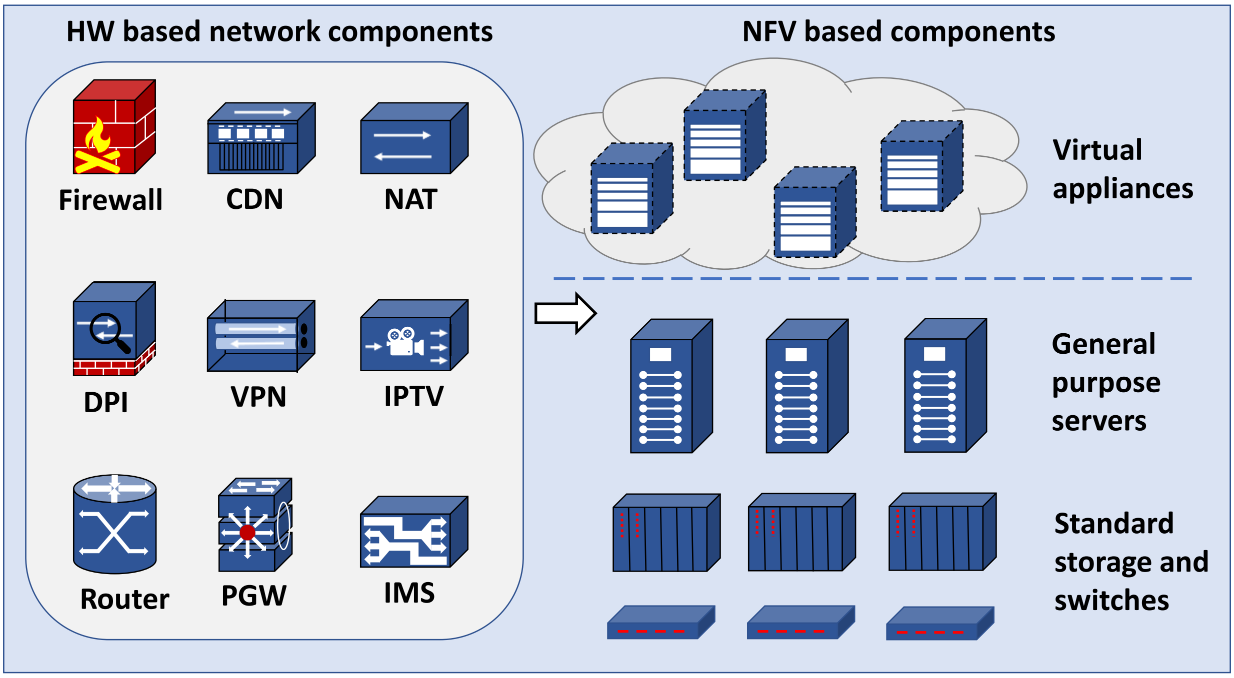


Figure 7. Dedicated HW and NFV based network components.

Common components of dedicated HW based approach presented in Figure 7 includes firewalls, Content Delivery Networks (CDN), Network Address Translation (NAT), Deep Packet Inspection (DPI), Virtual Private Networks (VPN), Internet Protocol Television (IPTV), routers, Packet Data Network Gateways (PGW) and IP Multimedia Subsystems (IMS). In NFV based approach same functionality is implemented as virtual SW components on general purpose servers together with standard storages and switches. [22]

NFV based architecture has three major differences compared to network built using dedicated components [22]:

1. Separation of software from hardware
    - Hardware and software evolve independent from each other
2. Flexible deployment of network functions
    - Network function SW can be automatically deployed and run in a pool of HW resources. These HW resources can be spread around several data centres and different SW functions can be run in any of those centres at any time.
3. Dynamic service provisioning
    - NFV performance is dynamically and flexibly scalable. As the need for performance increases network operators can then add more performance to the system. This way there is no underutilized HW meaning that investments are in efficient use.

NFV architecture framework is presented in Figure 8 below [22][23].

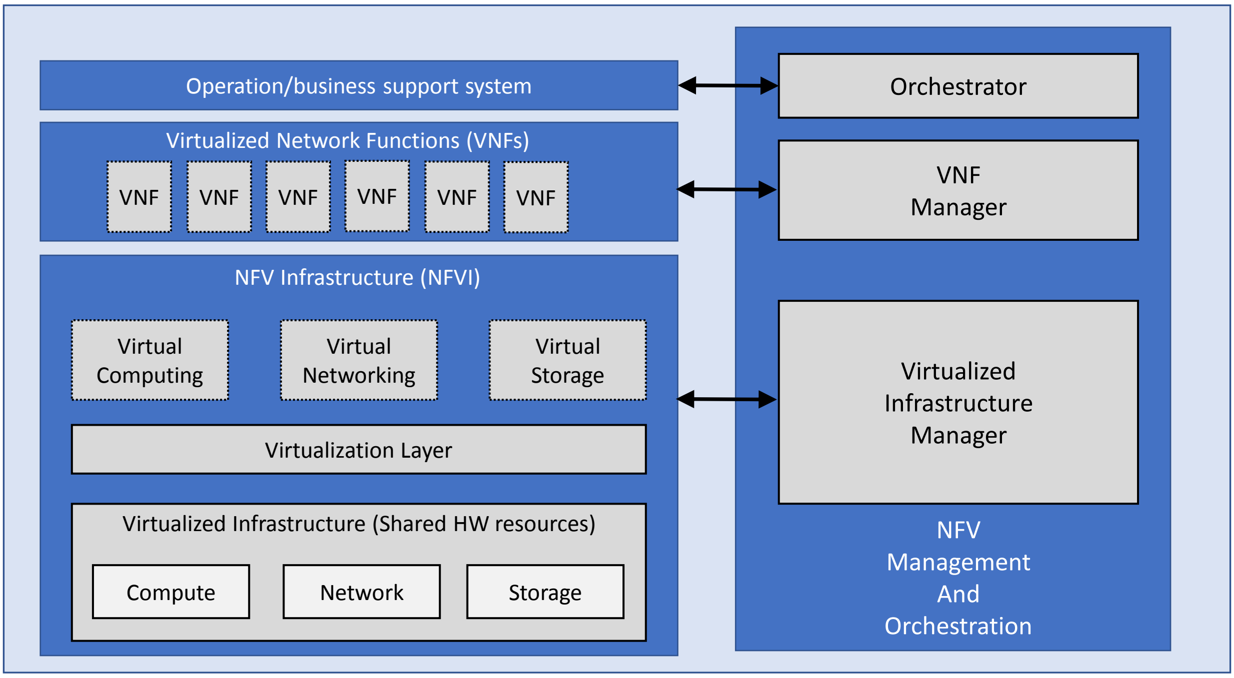


Figure 8. NFV architectural framework.

There are four main blocks in the NFV architecture [22]:

1. Orchestrator
    - Manages and orchestrates SW and virtualized HW resources.
2. VNF manager
    - VNF lifecycle management with instantiation, scaling, termination and updating of the VNFs.
3. Virtualization layer
    - Offers standardized interfaces to SW. Abstracts physical resources and ties VNFs to the virtualized shared HW resources. Offers Virtual Machines (VMs) and hypervisors for them.
4. Virtualized infrastructure manager
    - Virtualizes and manages computing, networking and storage resources. Control the interaction of those resources with VNFs.

From SDN and NFV architecture descriptions above it is clearly visible that technologies previously developed for cloud computing (like open switches, hypervisors, etc) and Commercial Off-the-Self Servers are essential for NFV. Interestingly those are also the sources for the weaknesses, challenges and threats NFV has. Virtualization may lead to network performance issues like latency variations and throughput instability. NFV networks needs to be able to co-exist and work with already implemented previous generation networks. Operators do need to ensure that network utilizing NFV does meet the same performance and stability that is reached by traditional networks. [22]

2.4.4 Virtual Network Function

Virtual Network Function is one of the main new items in 5G architecture compared to previous generations as the 5G core network is being virtualized. This technology has already been taken into use in cloud-based network architectures, so experience of the matter exists. In VNF functions that previously was run in dedicated HW is in VNF case run in commonly available servers as SW. [25]

As network functions that previously was run in HW are virtualized new kind of challenges emerge. VNFs can be dynamically placed in the network and amount of them can be dynamically increased or decreased. Therefore, the placement and amount of different VNFs is of the utmost importance. Operators can dynamically create several types of 5G networks by using varying amount of VNFs with different functionalities. So, the composition of the network can change based on the current demand of the traffic, varying number of customers or new slices that needs to be created. Operators are then faced with new kinds of problems that were not faced with previous mobile networks: how to effectively orchestrate 5G network to form a cost effective and performance optimized network with Quality of Service (QoS) that has been promised. [26]

2.5 Main modules of 5G network and 5G Test Network

In this chapter, the main modules of the 5G network will be presented and their functionality and purpose are shortly explained. Focus is kept on modules that are important from the perspective of this thesis so all the different modules that 5G has or supports will not be gone through.

2.5.1 Evolved Packet Core

EPC stands for Evolved Packet Core; the core network of the LTE system and it was first introduced in 3GPP Release 8. Unlike the previous generations that had all or part of their speech and data connected using circuit switching technologies, 4G system is from the beginning design to be all-IP, including speech. [27]

EPC has a flat architecture meaning that data traffic is handled performance and cost effective. There are only few network nodes that handle the transferred data and there is no need for protocol translation. User data and control plane are separated from each other so that scaling is made independent leading to ease of dimensioning and adaptation of the operator network. [27]

Figure 9 below presents the basic EPC architecture in 4G network [27]:

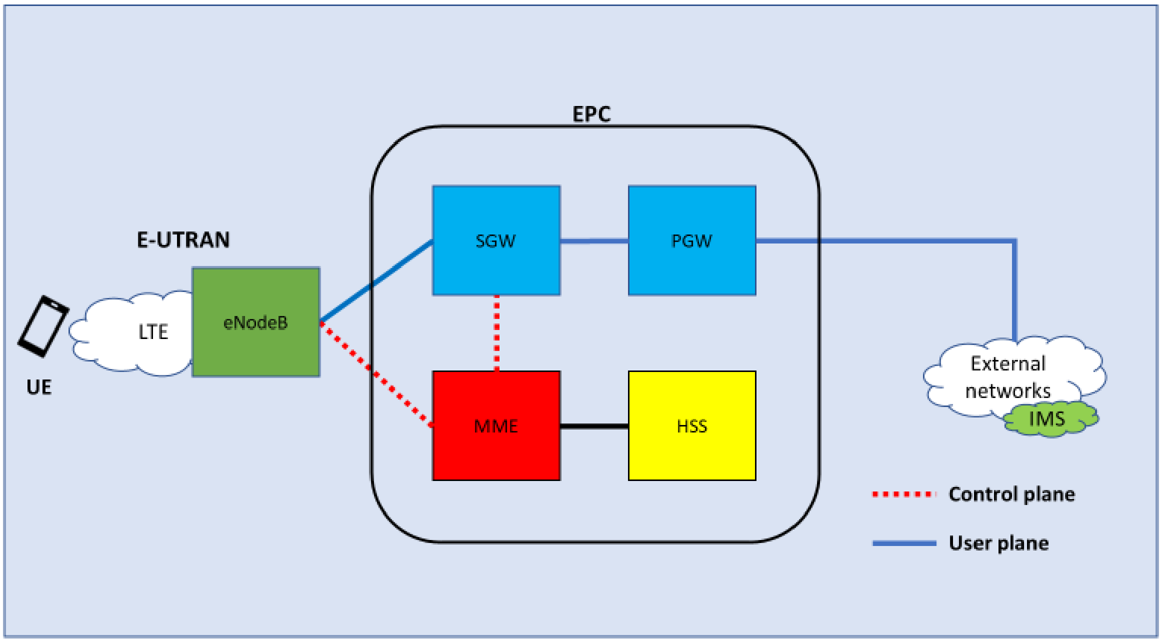


Figure 9. Basic EPC architecture in 4G.

In Figure 9 above User Equipment is connected to the 4G network over Evolved Universal Terrestrial Radio Access Network (E-UTRAN). The Evolved NodeB is the LTE base station connected to LTE radio. EPC is built from four elements: Serving Gateway (SGW), PDN Gateway, Mobility Management Entity (MME) and the Home Subscriber Server HSS. EPC is then connected to external networks. One of the external networks can be IP Multimedia Core Network subsystem IMS. Basic functionality of the EPC elements is explained below: [27]

Home Subscriber Server (HSS)

- database for user and subscriber related information like authorization, location etc. [28]
- delivers support functions in
    - mobility management
    - call and session setup
    - user authentication and access authorization

Serving Gateway (SGW)

- handles user plane
- connects User Equipment UE and EPC to each other
- transports incoming and outgoing IP packets
- anchor point for handovers (both intra and inter LTE)
- connected to PGW in EPC

Packet Data Network Gateway i.e. PDN Gateway (PGW)

- handles user plane
- connects EPC and external IP networks
- routes packets to and from the PDNs
- IP address and IP prefix allocation or policy control and charging

Mobility Management Entity (MME)

- handles control plane
- signalling related to mobility and security for E-UTRAN access
- responsible for tracking and paging of the UE in idle mode
- termination point for Non-Access Stratum (NAS) protocols

Unlike in the Figure 9 and as described in chapter 2.5.1 EPC can be connected not only to 4G but also to 5G new radio in the NSA mode that, e.g. 5G Test Network (5GTN) is using. That way already existing 4G infrastructure can be used as starting point for the 5G network. Operators' 4G investments and technical know-how can then be used in establishing 5G networks and ease the way from 4G towards the 5G.

2.5.2 5G Test Network

This thesis and the related tests are done utilizing 5G Test Network built to Oulu University campus and governed by the University of Oulu. Figure 10 below shows simplified picture of the 5GTN system.

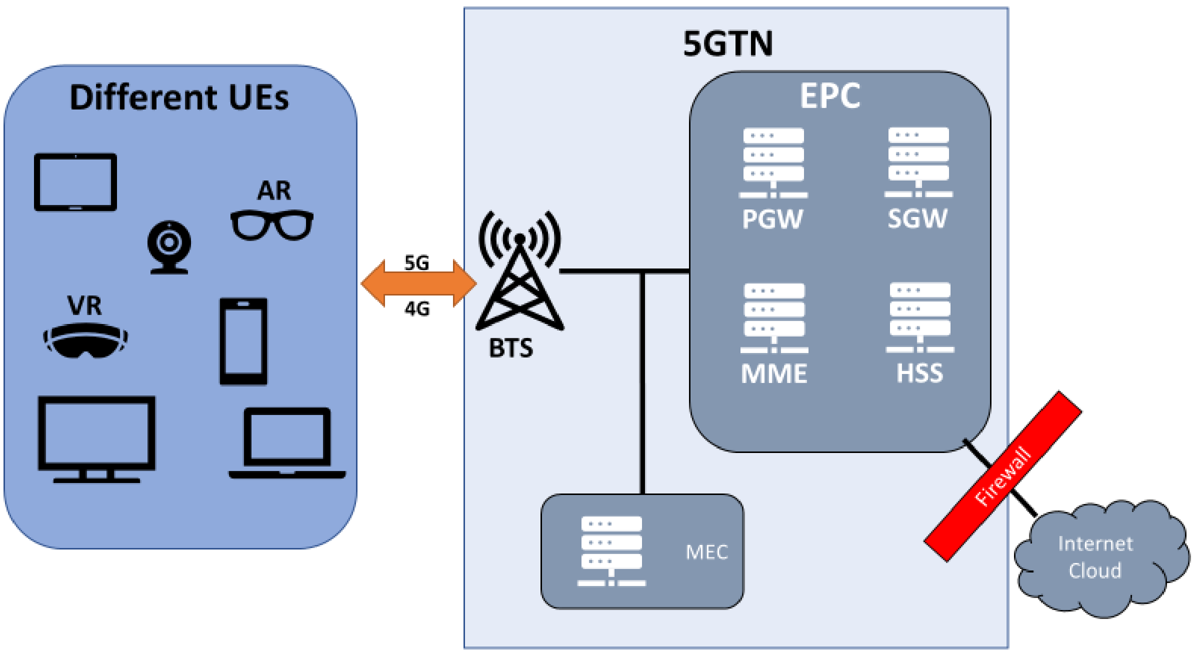


Figure 10. Simplified 5GTN architecture.

In Figure 10 the 5GTN has been divided into two parts: network equipment and the User Equipment's. Major parts of the 5GTN system are different end user equipment that can be used to access the network like mobile phones, computers, AR and VR glasses, pads, cameras etc. On the network side Base Transceiver Station (BTS), MEC and Evolved Packet Core (EPC) that is built from PGW, Serving Gateway (SGW), Mobility Management Entity and Home Subscriber Server are the main building blocks. 5G network connection to internet is protected by firewall.

5GTN works in a Non-standalone NSA mode. There are several options for Standalone SA and non-standalone implementations and those are presented in Figure 11 below [29].

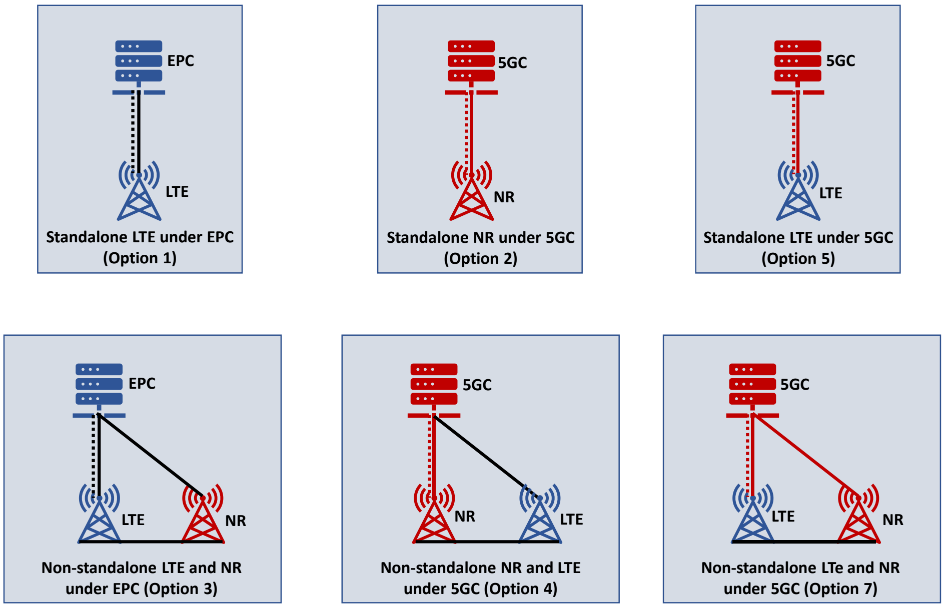


Figure 11. Different NSA and SA options.

5GTN is based on Option 1 and Option 3, where 5G radio is added to already existing 4G network allowing dual connectivity where compatible devices can utilize both LTE and New Radio access. This option uses already existing 4G network and EPC as basis and adds 5G NR to it. This accelerates 5G time to market for operators and also reduces the investment costs. In other options there are multiple combinations of EPC, 5G Core (5GC) and NR to introduce different ways to launch 5G service. [29]

2.5.3 5G New Radio

Fifth Generation mobile phone network radio interface New Radio is specified by the Third Generation Partnership Project 3GPP. Targets for the specification are set by 5G use cases, like MBB, IoT or URLLC. This leads to large set of diverse targets like high system capacity, large coverage, high data rates, low device cost, low energy consumption, ultra-high reliability and low latency. [30][31]

NR extends the used radio spectrum significantly compared to previous generations. Used frequencies goes from below 1 Giga Hertz (GHz) all the way to 52.6 GHz. With Millimetre-Wave (mmWave) frequencies, very large channel bandwidth can be used and due to that very high data rates per channel are enabled. With high frequencies radio signal attenuation is high and therefore network coverage is limited compared to previous generations operating at lower frequencies. This can be compensated by usage of multi-antenna technologies like MIMO, but only partially. Similarly, Ultra Dense Networks (UDN) and Cloud RAN (CRAN) technologies can be used to implement highly efficient ultra-dense networks. [30][31]

2.5.4 Multi-access Edge Computing

Multi-Access Edge Computing is one of the main new features 5G brings. Purpose of the MEC is to offer a computing platform close to the radio interface and therefore close to end user. That way software services that previously could be placed only to the internet cloud can now be brought close to the user. There is a lot of use cases that might profit from the possibility to use MEC platform.

MEC as part of the 5G system is still being specified and there are different definitions of the MEC still around. For some MEC is just a server located at the edge of the network where whatever application can be run. For some it is more than that. In this definition it is a specified virtualized environment that is an integrated part of the 5G network. In this thesis focus is in this larger definition.

MEC is one of the main technologies used to reach the demanding targets set by Key Performance Indicators (KPIs) of the 5G networks, especially low latency and bandwidth efficiency targets. Edge computing enables the application to be run in a server at the network edge near the base station instead of the internet cloud server. This way application hosting is brought close to the end user and data generated by mobile applications or for example IoT sensors. As mobile phone network is turning into service platform through technologies like SDN and NFV also MEC plays a major role in it. Edge computing opens the network edge for software applications and services not only for operators but also for third parties. MEC enables also the possibility to explore new use cases that have strict latency and bandwidth efficiency requirements. [32][33]

There are several implementation options for MEC. MEC functionality can be installed to Edge server as a native Linux application and, in that case, it is usually called just MEC. If MEC is implemented as virtual application on top of the NFV virtualization infrastructure into the Edge server it is then often called virtual MEC (vMEC).

MEC host runs applications as software-only entities. Applications are run on top of Virtualization infrastructure like NFVI that is located in the MEC host. MEC framework is presented in Figure 12 below. [34]

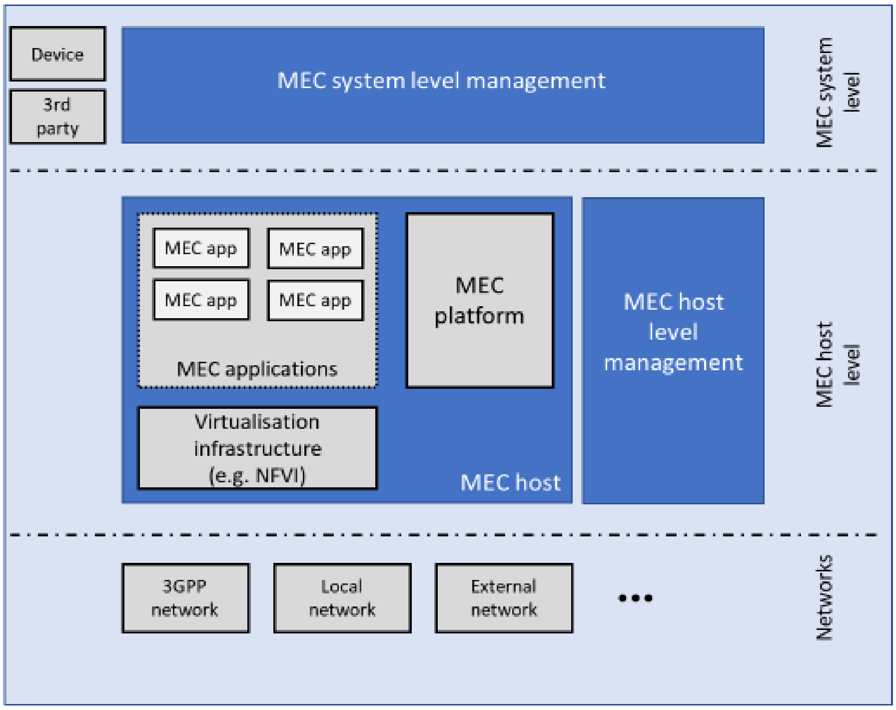


Figure 12. Multi-access Edge Computing framework.

Multi-access Edge Computing that was presented in Figure 12 has following entities [34]:

- MEC host
  - MEC platform
  - MEC applications
  - Virtualization infrastructure
- MEC system level management
- MEC host level management
- external related entities, i.e. network level entities.

MEC host and MEC host level management are necessary to run the MEC applications. MEC host contains a MEC platform and virtualization infrastructure to offer computing, storage and network resources for the MEC applications. MEC platform has the functionality to run MEC applications on specific virtualization environment and it can also provide services. MEC applications are run in the MEC host virtualization environment as governed by MEC management. [34]

MEC management is made of MEC system level management and the MEC host level management. MEC system level management core component is Multi-access edge orchestrator that has an overview of the complete MEC system. MEC host level management includes the MEC platform manager and the virtualization infrastructure manager. It handles both the management of MEC specific functionality of a MEC host and the applications that are running on it. [34]

MEC is meant not only for 5G networks but it can be deployed also in 4G network as its architecture is agnostic to weather the network is 4G or 5G. As 4G is the dominant mobile phone network architecture for the coming years it is natural that operators are first working towards running MEC in their existing 4G networks. That is also the way to go for the operators that are building their network based on the non-standalone architecture where 5G

NR is added to already existing 4G network. MEC naturally needs to work with both 4G and 5G radios in NSA network. [33]

Main function of the MEC platform is to route IP packets to different applications running in the MEC. These applications can handle the data traffic routed to them in four different ways [33]:

- Breakout mode: session connection is redirected to a MEC application that is hosted locally or on a remote server. This mode is typically used for Content Delivery Network, gaming, media content and enterprise Local Area Network (LAN).
- In-line mode: In this mode session connectivity is kept with the internet server and all traffic is just passing through the MEC application. Used for example for transparent content caching and security applications.
- Tap mode: Here specified traffic is duplicated and then forwarded to the tap MEC application. Can be used for example when deploying virtual network probes or security applications.
- Independent mode: No traffic offloading function is needed. MEC application is still registered in the MEC platform and it will receive MEC services like DNS and Radio Network Information Service (RNIS).

It is possible to deploy MEC into the network in multiple ways. Even though many choices are possible, basic scenarios how MEC can be located are presented below. Both 4G and 5G scenarios are presented.

From deployment options presented below two main categories can be seen: one that has MEC as a standalone module and the other where one or more EPC functionalities are collocated with MEC [33]. All of these have their benefits and down sides and it is the task of the operator to select the deployment option(s) that best fit their needs.

Bump in the Wire (4G)

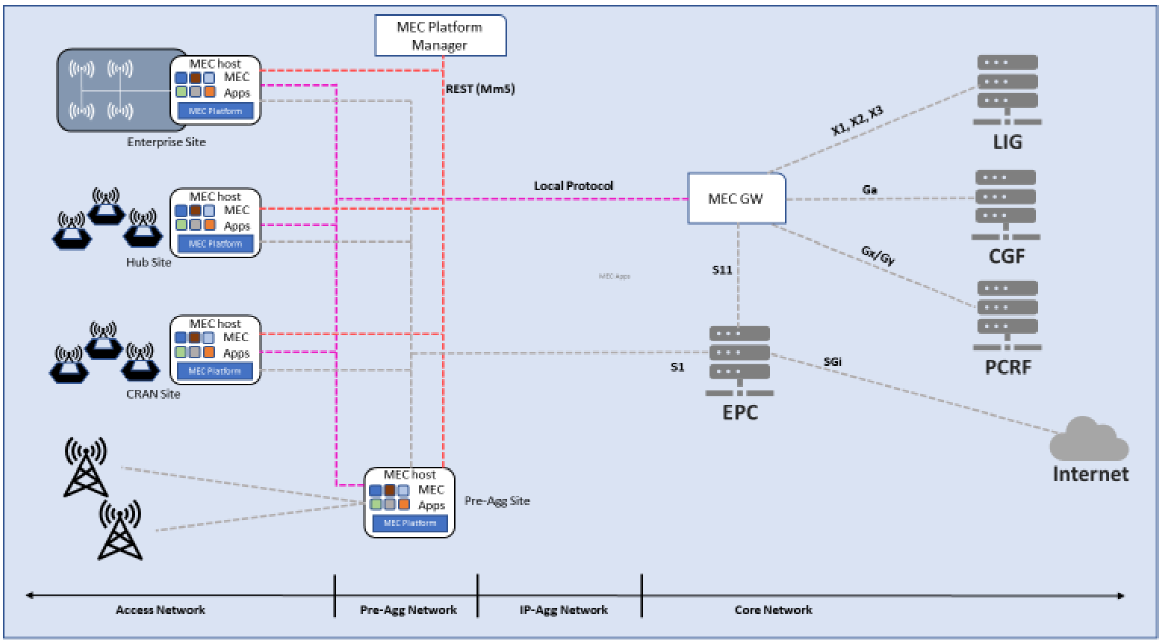


Figure 13. MEC deployment "Bump in the wire".

In Figure 13 LIG stands for Lawful Intercept Group, CGF means Charging Gateway Function and PCRF is abbreviation for Policy and Charging Rules Function. This scenario covers all possible scenarios where MEC platform is installed between base station and the mobile core network. Here MEC platform routes plain IP packets to and from MEC applications (local switching mode) and can also route General Packet Radio Service (GPRS) Tunnelling Protocol -encapsulated (GTP-encapsulated) packets to and from the Serving Gateway for regular traffic (legacy S1-U mode) if MEC is bundled together with evolved Node B (eNB). This MEC location is usable for enterprises as it allows intranet traffic breakout to local services. [33]

In all other possible MEC locations like close to the radio node or at an aggregation point MEC is connected to 4G network via S1 interface. In these locations MEC host's data plane needs to process user traffic encapsulated GTP-User plane (GTP-U) packets. With this MEC location low latency is supported if MEC is located close to the eNB or any other location that has minimal latency. [33]

Distributed EPC (4G)

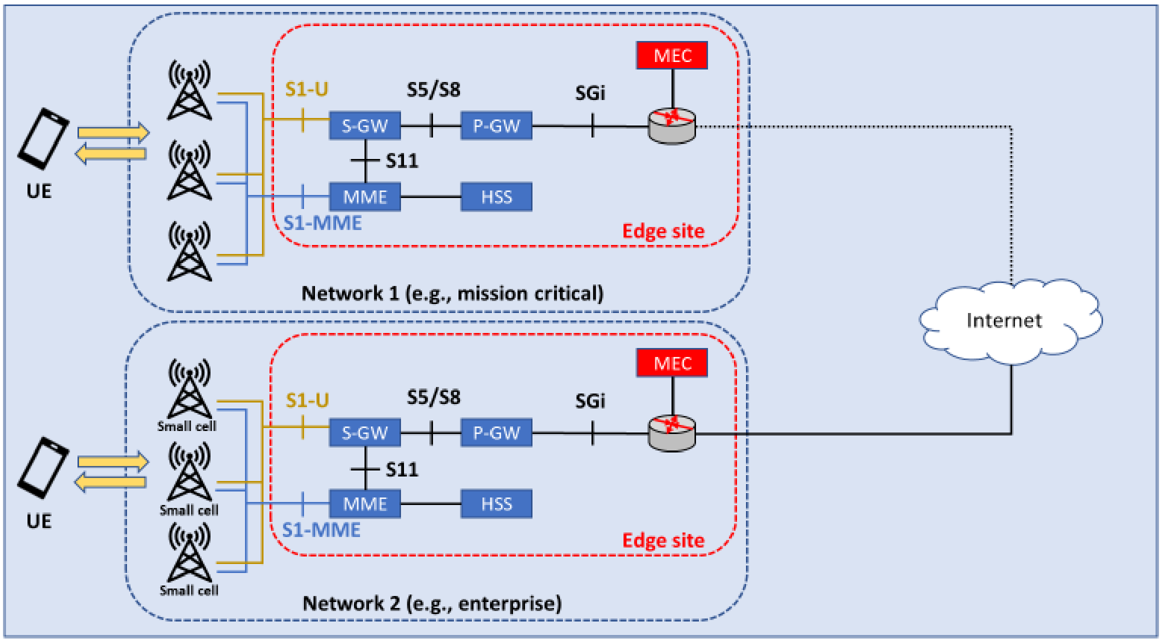


Figure 14. MEC deployment with distributed EPC.

As the name suggests in this scenario MEC host includes all or part of the EPC. MEC data plane is connected to SGi interface. Mission Critical Push to Talk (MCPTT) and M2M communications where it is optional to communicate with the operator's core site benefits from this kind of set-up. Here also HSS is co-located with MEC and therefore local services can run even without a working backhaul. This deployment scenario can be used for example by first responders, public safety and mission critical industrial sites. Scenario is presented in Figure 14 upper part. [33]

In some cases, HSS is unique and it is centrally managed by the operator at the core site. Also, operator's core site PGW can be used for some selected Access Point Names (APN's) and this enables local management of the subscriber database and the local EPC can be used in the MEC to offload the entire APN traffic. Distributed EPC enables the possibility to deliver the QoS and configurability that for example enterprise customers require. This scenario is presented in the Figure 14 lower part. [33]

Note that in Figure 14 the systems look the same. There are two different systems for two different needs: mission critical and enterprise.

It is possible to co-locate EPC functions and the MEC applications in the same MEC host. EPC and MEC components can run as VNFs on the same NFV platform. Therefore, operators can better utilize resources, resulting in improved scalability of the network resources and direct cost savings. Example of this deployment scenario is in Figure 15. [33]

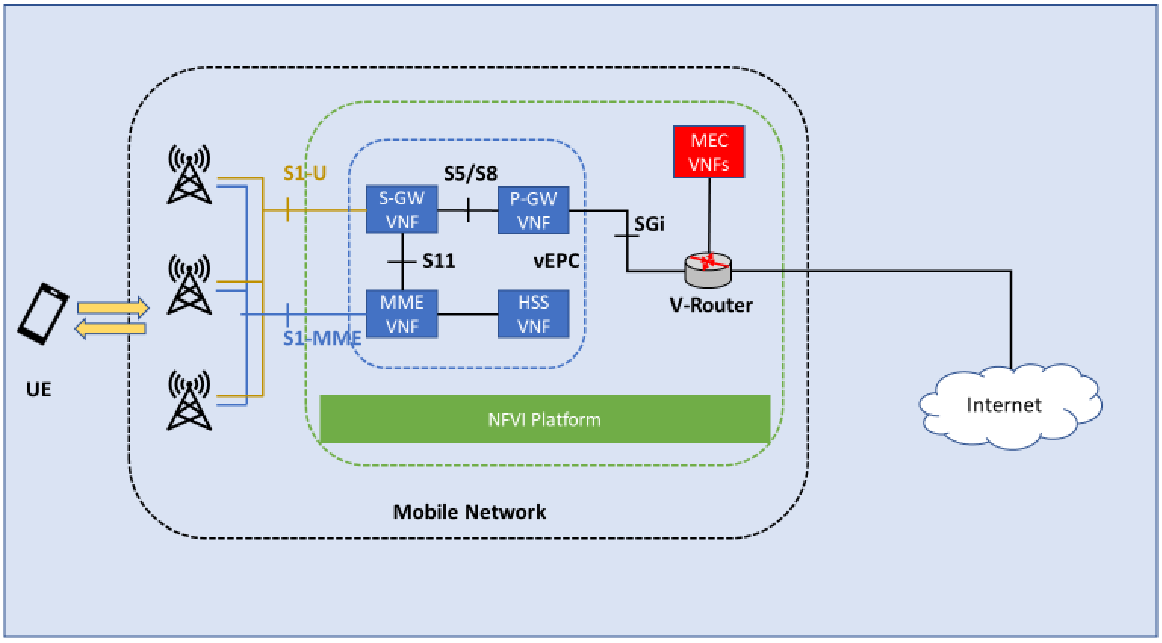


Figure 15. MEC deployment with MEC and EPC on the same NFVI platform.

Distributed S/PGW (4G)

In distributed S/PGW deployment option only SGW and PGW are deployed at the edge site. Other EPC functions (MME and HSS) remain at the operator's network core site. Also, here MEC is connected to PGW through SGi interface. With this kind of deployment operator has full control over the MME function. This option is illustrated in Figure 16 below. [33]

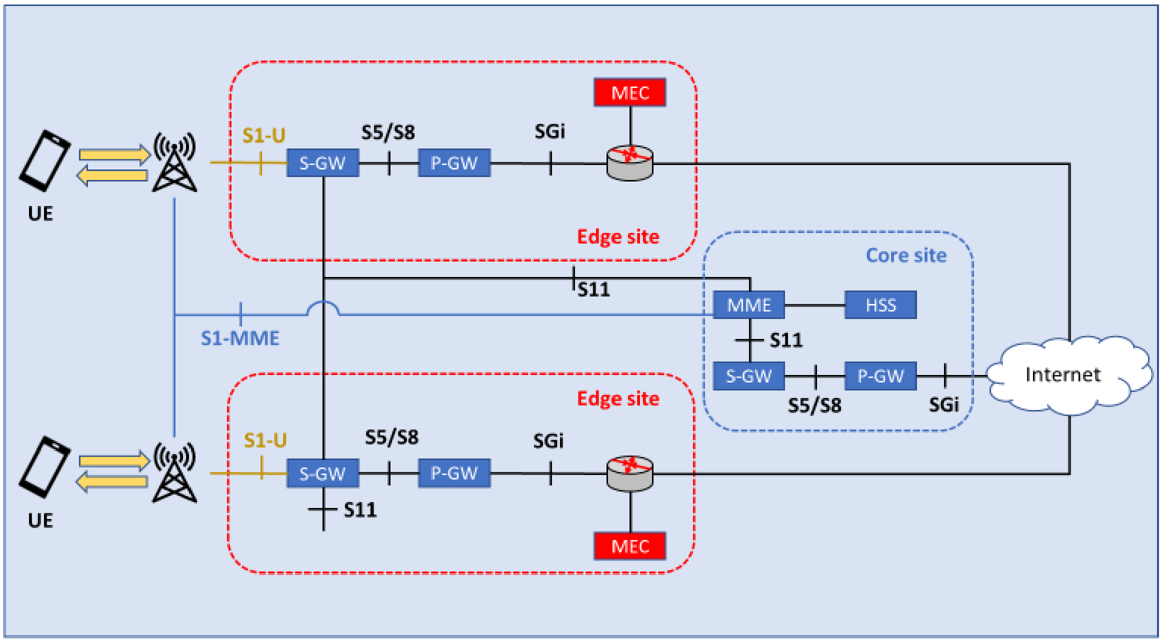


Figure 16. MEC deployment "Distributed S/PGW".

In this option both distributed EPC components PGW and SGW can be run as VNFs together with MEC application on the same MEC host just as in the previous option.

Distributed SGW with Local Breakout (SGW-LBO) (4G)

Operators' need to have greater control over the granularity of the data that is steered has led to this deployment option. Here end users have a need to reach both operator's core network and MEC applications in a selective manner over the same Access Point Name. In this option MEC can be deployed on the same edge site or even on the same MEC host. Both MEC and SGW can even be hosted as VNFs in the same MEC platform. This option is described in Figure 17. [33]

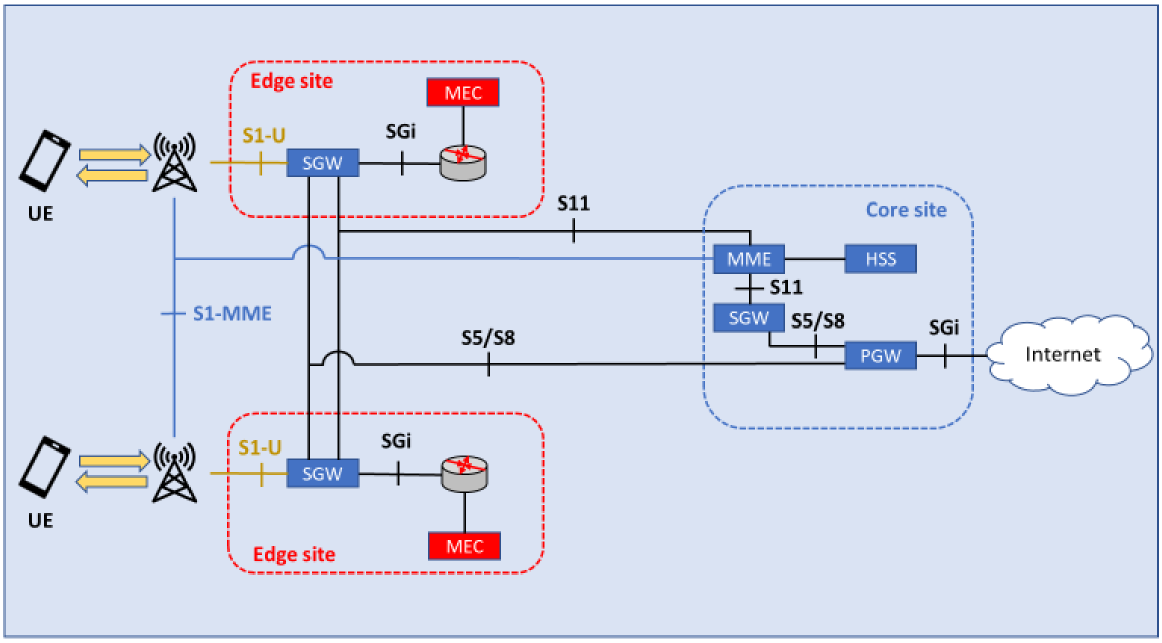


Figure 17. MEC deployment "Distributed SGW with Local Breakout (SGW-LBO)".

MEC deployment in the 5G architecture network

5G network has Service Based Architecture (SBA) that contains multiple control plane functional entities. These are Policy Control Function (PCF), Session Management Function (SMF), Application Function (AF), etc. There are also data plane functional entities like User Plane Function (UPF). As a change from previous generation mobile phone networks 5G allows more flexible data plane deployment with better support for edge computing. Therefore, MEC can be easily integrated into a 5G network. Figure 18 presents one way of connecting MEC into the 5G network. [33]

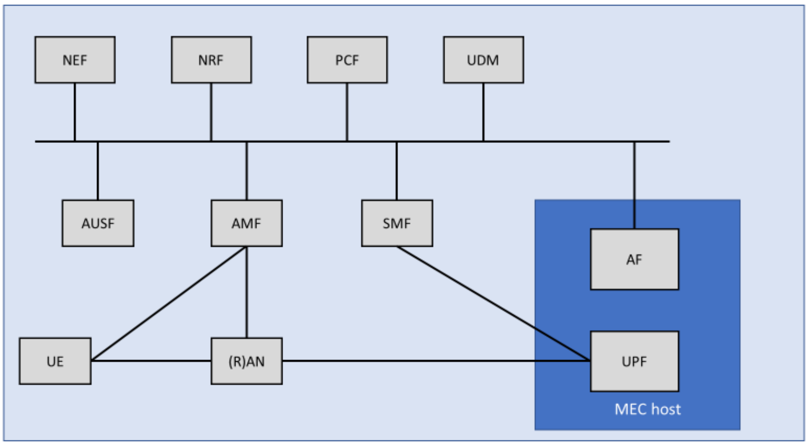


Figure 18. MEC mapping to 5G network.

In Figure 18 Network Exposure Function (NEF) secures provision of information from external application to 5G network, Network Function (NF) Repository Function (NRF) handles service discovery function and maintains NF profile and also available NF instances, PCF among other things covers unified policy framework and provides rules the CP function and Unified Data Management (UDM). Authentication Server Function (AUSF) as name suggests handles authentication server functionality. Access and Mobility Management Function (AMF) handles termination of Network-Attached Storage (NAS) signalling, mobility management, access authentication and authorization, etc. Session Management Function covers not only session management operations like session establishment, but also many other functions like Dynamic Host Configuration Protocol (DHCP) functions, UE IP address allocation, etc. Application Function AF covers among other things application influence on traffic routing and accessing NEF. Lastly User Plane function UPF covers functions like packet routing and forwarding, packet inspection and QoS handling. [35]

In the example presented in Figure 18 the MEC platform performs the traffic routing and steering function in the UPF. PCF and SMF can set a policy to influence the traffic routing in the UPF. As AF is connected to PCF it can influence the traffic routing and steering. Therefore, MEC is capable of influencing the UPF through the standardized control plane of the 5G.

As with 4G there are several MEC deployment options in 5G also. Figure 19 shows how to migrate from different 4G MEC deployment options to similar 5G deployment options. Here the functionalities are preserved but the implementation is according to 5G system architecture. [33]

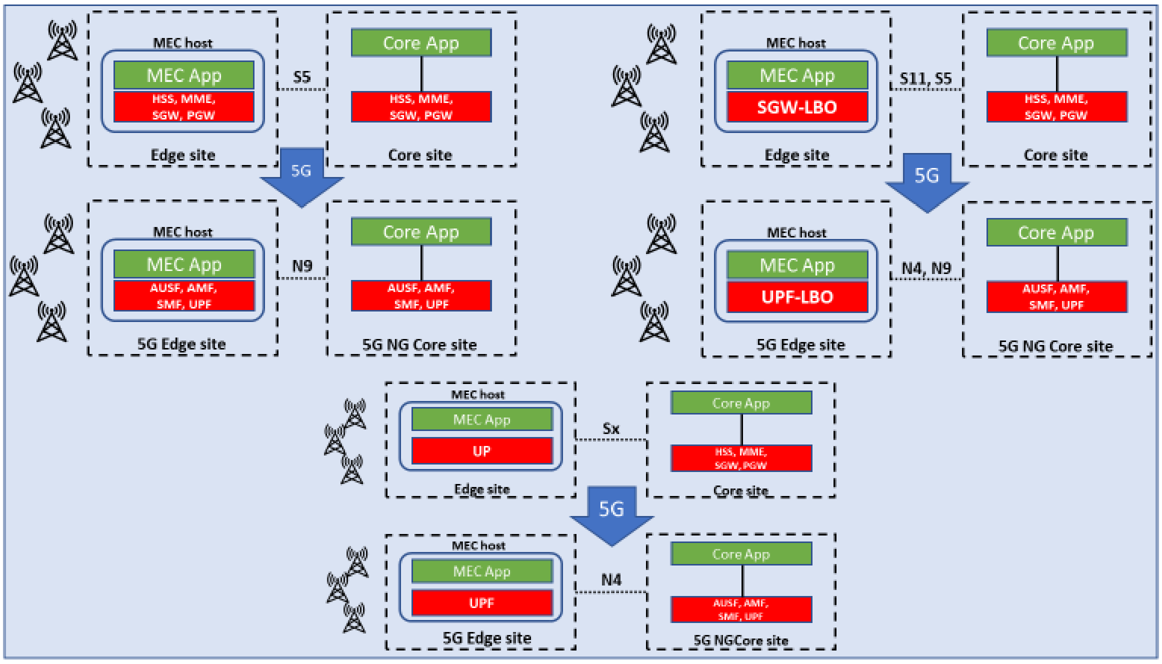


Figure 19. MEC deployment migration patterns from 4G to 5G.

Figure 19 shows migration from 4G to 5G for the following deployment options: top left corner shows MEC deployment with distributed EPC, top right corner presents MEC deployment distributed SGW with local breakout and low part of the figure represents the MEC deployment with Control/User Plane Separation (CUPS) option. [33] There is no one deployment option that is best but as with 4G networks, operators are to choose the deployment option that best suit their needs in the 5G environment.

2.6 On 5G Performance

2.6.1 5G performance improvement promise against 4G

There are expectations how 5G will bring improvements over the 4G networks. Some of these expectations are documented in International Telecommunication Union Radiocommunication Sector (ITU-R) document “IMT vision – Framework and overall objectives of the future development of IMT for 2020 and beyond”. International Telecommunication Union (ITU) document “Key features and requirements of 5G/IMT-2020 networks” then summarises the high-level targets for 5G. Based on these documents Figure 20 shows how the key capabilities are expected to change going from 4G towards the 5G: [36][37]

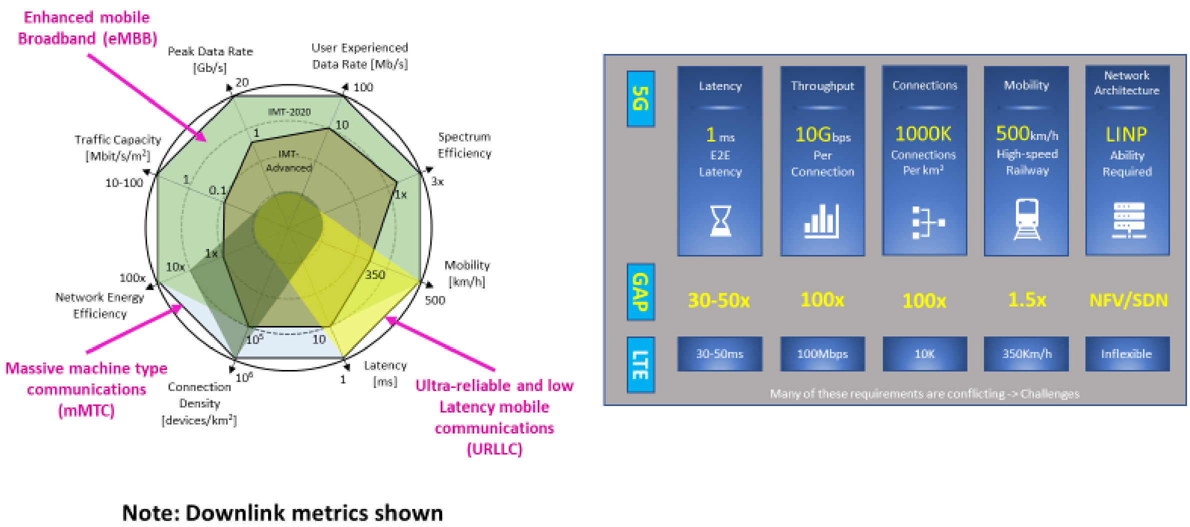


Figure 20. Performance targets for 5G.

Main improvement targets for 5G as can be seen in Figure 20, are:

- latency from min 10 milliseconds (ms) to 1 ms
- channel throughput from 1 gigabit per second (Gbps) to 20 Gbps
- end user data rate from 10 megabits per second (Mbps) to 100 Mbps
- supporting moving vehicles from 350 kilometres per hour (km/h) to 500 km/h
- from 10 simultaneous connections to 1000 simultaneous connections

As some of the requirements are contradicting and not all of them can be achieved simultaneously for single use case. It is the operator network implementation that in the end will determine the final performance of the 5G networks. Due to that different networks will have varying performance and there is not one set of performance figures that do match for all the 5G networks. That can be seen as a benefit, since operators can tune their network for their own and their customer’s needs.

2.6.2 Key Performance Indicators of 5G

Key Performance Indicator KPI is a measurable value that is used to evaluate how well targets are met. In 5G network environment KPIs set the target values for defined measurable technical features or requirements. KPIs define not only the value but also the condition under what the value is to be reached. There can be different KPI requirements for different categories for one requirement type. Table 2 lists the main KPIs set for the 5G network according to 5G Infrastructure Public Private Partnership (5G PPP). [38][39][40]

Table 2. Main Key Performance Indicators for 5G

5G performance requirement type	Minimum KPI requirement	Category
Peak data rate	Downlink: 20 Gbps Uplink: 10 Gbps	eMBB
Peak spectral efficiency	Downlink: 30 bits/sec/Hz Uplink: 15 bits/sec/Hz	eMBB
Data rate experienced by the user	Downlink: 100 Mbps Uplink: 50 Mbps	eMBB
Area Traffic Capacity	Downlink: 10 Mbits/sec/m^2 in indoor hotspot (eMBB test environment)	eMBB
User Plane latency	4 ms for eMBB 1 ms for URLLC	eMBB, URLLC
Control Plane latency	10 ms	
Mobility	up to 500 km/h	eMBB
Mobility Interruption Time	0 ms	eMBB, URLLC
Reliability	A general URLLC reliability requirement for one transmission of a packet is $1*10^5$ for 32 bytes with a user plane latency of 1ms	URLLC
Connection density	1 million devices / km^2	mMTC

Main part of the KPIs are for enhanced Mobile BroadBand eMBB category but also Ultra-Reliable and Low Latency Communications URLLC plays a large role. In this thesis the focus is set on KPI User Plane Latency. As seen the target is one millisecond for URLLC and 4 milliseconds for eMBB category.

3 HOSPITAL USE CASE PERFORMANCE

The focus of this thesis is on 5G delay measurements of a demo that is named as hospital use case demo. The demo itself will be presented in the chapter 3.1. The chapter 3.2 presents different implementation options how the system can be built. Chapters 3.3 and 3.4 clarify the system setup for Edge server or internet cloud implementations.

3.1 Hospital use case

Hospital use case demo is developed in the University of Oulu to demonstrate new possibilities that 5G networks will offer. It was first introduced in the 6G summit in Levi April 2019 and since then it has been further developed and used in several seminars and other sessions.

In the demo a situation where a patient is being either treated or operated is video streamed to an observation room using 5G network. There is a 360° camera in the operation room that will give the overall view (to all directions, i.e. 360 degrees) of the room and the live video is streamed to audience in the observation room. 360° stream can be viewed in the observation room using VR glasses. One or more medical personnel in the operation room wear AR glasses that stream their live view to the observation room. Stream from the AR glasses cameras can be viewed using for example VR glasses, mobile phones, separate monitors or TVs. All streams are sent and received using 5G network.

There are sensors in the operation room that can measure the environment and equipment and sensors attached to the patient measuring relevant data like blood pressure, heart beat rate, etc. Medical personnel that wears the AR glasses do also get sensor feed to the AR glasses display. Sensors and the data shown can be selected separately for each of the glasses.

Figure 21 below describes concept for the hospital use case demo.

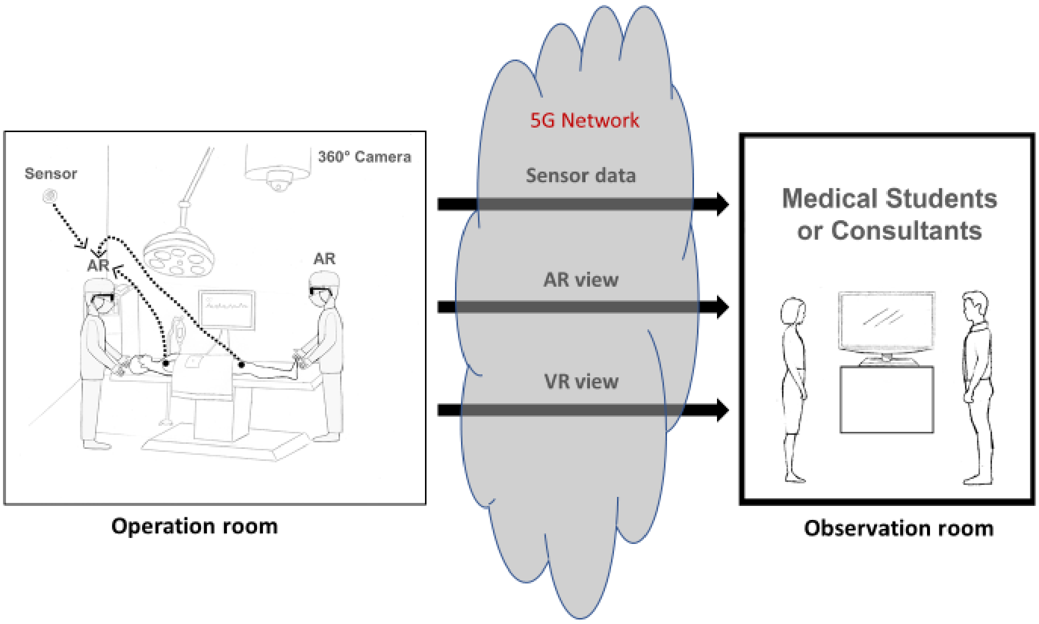


Figure 21. Hospital use case demo concept.

This use case was developed for the medical needs. It can be utilized for several different cases varying from home hospital and remote consultation to education. Few of these are listed below.

- **Remote consultation**
  - There is an urgent need for treatment, but the required specialist is not available as quickly as needed. With this setup specialist can consult remotely as he/she has a live real time view of the operation.
- **Home hospital**
  - Nurse that wears the AR glasses can treat the patient at patients' home. Doctor can then follow the procedure and instruct the nurse/several nurses as needed.
- **Education**
  - Students can follow the operation from several different angles in real time. Students can participate the training remotely from several different sites. Teaching is not tied to certain teaching facilities but can be flexibly organized from different types of situations and locations.

It is also notable that all the feeds can be recorded for further use both in education or in analysing the treatment and operations themselves.

3.2 Different implementation options

There are a few different options how the setup can be built. In this chapter those options are explained. The set is composed of the operation room with cameras, laptops, AR glasses and modems needed to send the video streams, 5G Test Network or internet cloud server with multimedia server capabilities and observation room with iPad, VR glasses and mobile phones that are needed to view the streams. Options differ mainly where the multimedia processing is done, in the 5G Test Network or the internet cloud server.

System uses NGINX Open Source multimedia server software for the stream protocol change and streaming server operations. NGINX processes all the streams and re-sends them to viewers. NGINX location and implementation method is the main differentiating item between different implementation options. First of all, it can reside in the 5G Test Network or in the internet cloud. NGINX in the internet cloud can be seen as the legacy way as that is how the NGINX had to be located before the introduction of the MEC.

Second location for the NGINX is the 5G Test Network. NGINX is installed on a server in the 5GTN. Multimedia server can be installed on its own or into the MEC. If it is installed as native Linux application, it is not registered as a 5G network service. NGINX cannot be accessed as a 5G MEC service but directly using the IP address of the server where it was installed.

If NGINX is installed as MEC application, it is registered to 5G network as a service and it is controlled by the 5G network. It also gets the intrusion and hack protection that 5G network and MEC can deliver. Multimedia server functionality is then offered as network service to the end users.

MEC as well as NGINX can be installed to the server environment as separate applications or as a virtual function. If MEC is installed as an application, NGINX can be installed as a separate application alongside MEC or into the MEC as container application or as virtual application. If MEC itself is installed as virtual application, then also NGINX is to be installed as virtual application alongside or into the vMEC.

Focus of this thesis is to test and measure the impact that different NGINX installation locations and methods have on its performance.

3.3 Utilizing MEC server for multimedia server SW

As described in the previous chapter, NGINX multimedia server software can be installed into the MEC server. There are several options how to install the NGINX:

- Native Linux application alone
- As a virtual application
- Into the MEC as a container application or virtual application

Basic architecture of the setup is described in Figure 22:

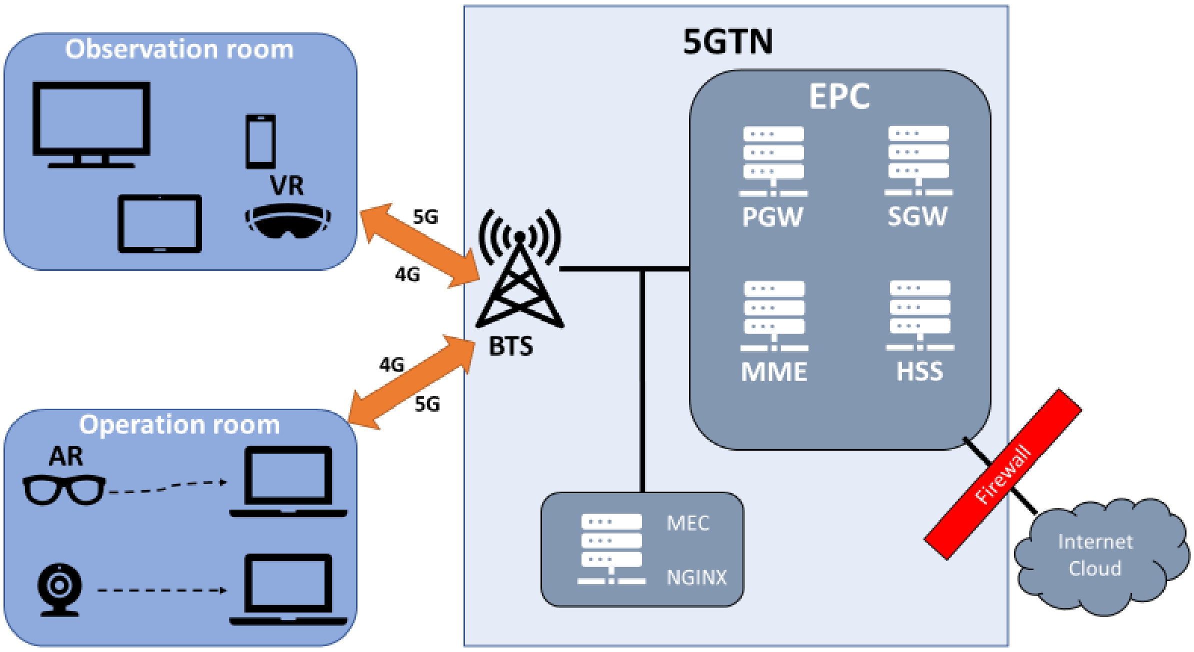


Figure 22. Hospital use case demo setup.

As seen in Figure 22 both the Operation room and the observation room equipment are connected to 5GTN. EPC is located on its own server in the 5GTN. NGINX is installed to its own Edge server separately or inside the MEC. Figure 23 shows how the streams flow in this kind of setup.

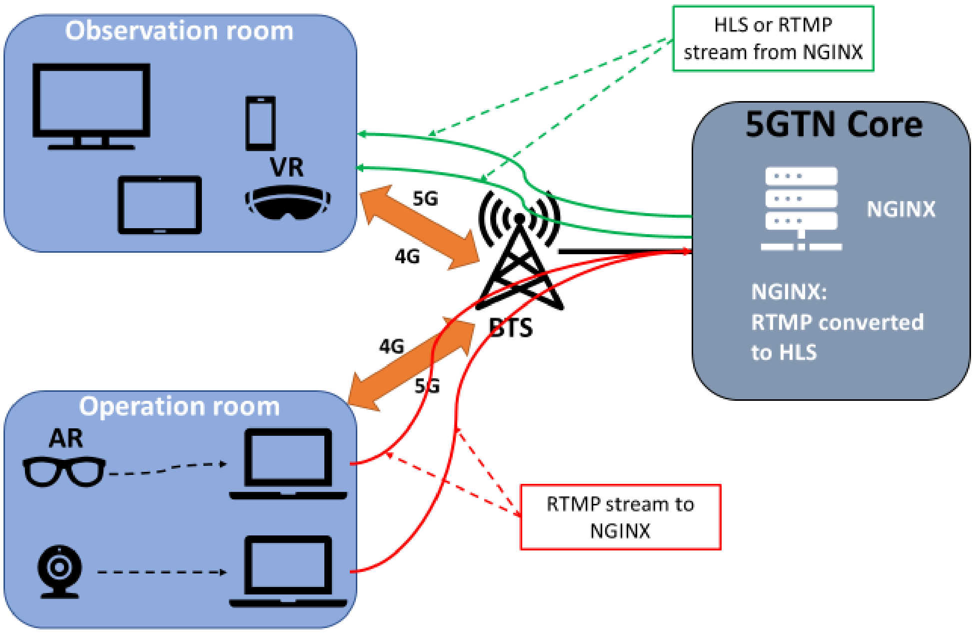


Figure 23. How different streams flow when NGINX in Edge server.

All streams from the operation room are sent as RealTime Messaging Protocol (RTMP) streams to NGINX multimedia server. NGINX does make the conversion from RTMP to HTTP (Hypertext Transfer Protocol) Live Stream (HLS). NGINX then acts as a multimedia server and hosts both HLS and RTMP streams to be viewed by streaming clients in the observation room.

In the 5GTN data packets are routed through the SGW and PGW to the NGINX, no matter if the NGINX is installed independently to Edge server, into the MEC or to the cloud.

3.4 Utilizing cloud for multimedia server SW

The legacy way of locating the NGINX multimedia server is to have it in the internet cloud. NGINX is installed as stand-alone application to a cloud server. RTMP stream is then routed to it, format conversion from RTMP to HLS is done and finally HLS and/or RTMP streams are transferred to observation room for viewing. system is described in Figure 24.

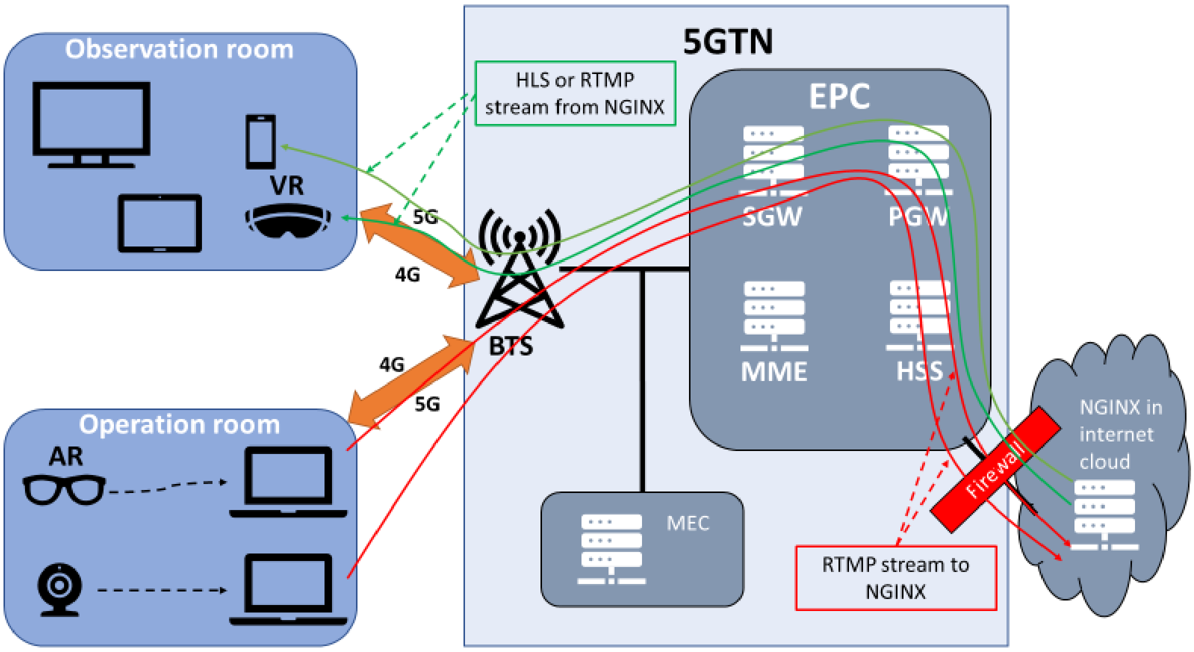


Figure 24. How different streams flow when NGINX in internet cloud.

Physical routing is expected to be much longer in this setup than in the one where NGINX is installed in Edge server in the 5GTN. Data packets are routed through SGW and PGW and through the firewall to the NGINX residing in an internet cloud server and vice versa when streamed to the observation room.

4 TEST ENVIRONMENT

4.1 Purpose of the tests

Purpose of this Diploma Thesis is to better understand the benefit of utilizing computing services at the Edge in the 5G mobile network instead of using internet cloud services. Intention is not to test all possible use cases or features of the Edge but to focus on delay.

One of the main promises of 5G is to decrease the user plane delay to millisecond level. More specifically on of the Key Performance Indicators of the 5G is to have a 4 ms delay target set for eMBB type of use cases and as small as 1 ms for URLLC.

Hospital use case is used as means to create a setup where the delay can be measured. Measurements are done as end to end measurements for sending the stream to NGINX multimedia server and also receiving the stream from it. That way both uplink and downlink measurements can be separated, and the results analysed both separately and from the whole system perspective. NGINX installation can be located at the Edge server either on top of Linux OS or into the MEC. Legacy method of locating the NGINX to the internet cloud services will also be tested and used as a reference.

4.2 Generic test setup view

Hospital use case's technical environment and the actual test environment are a bit different from each other as the test environment is simpler for practical reasons. These test environments and their differences are presented in this chapter. Figure 25 describes hospital use case environment.

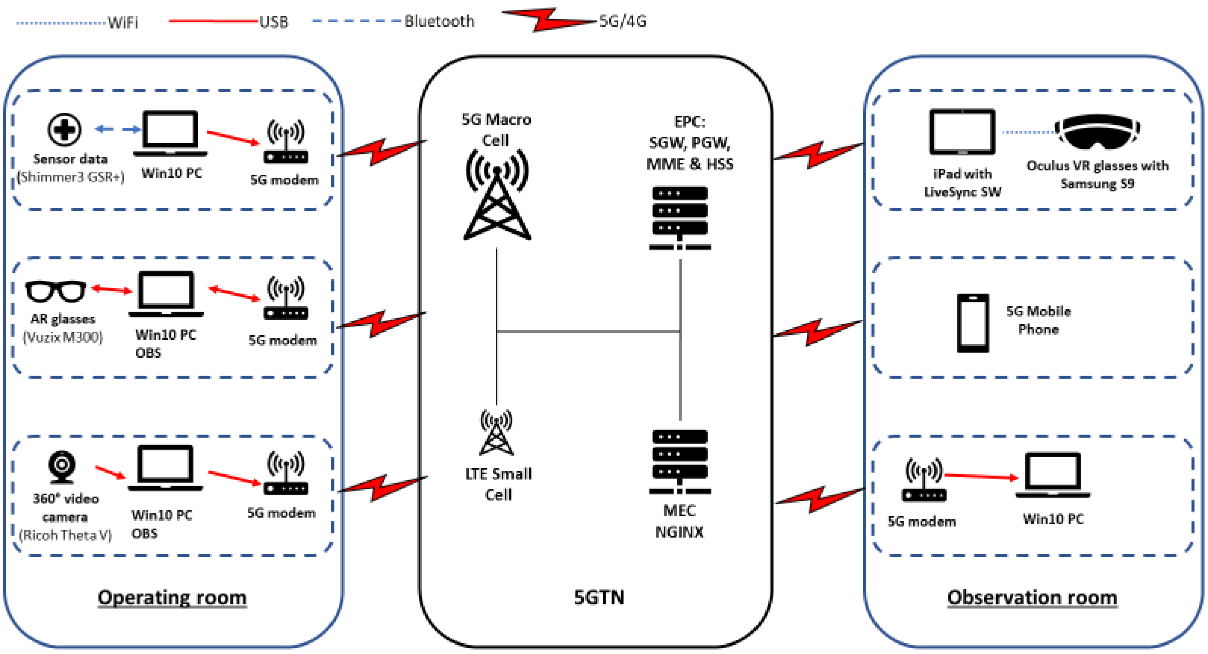


Figure 25. Hospital use case technical setup.

As can be seen from the Figure 25, there are two streams and the sensor data sent from the operating room. In the receiving side in the observation room there are three different devices receiving the streams: iPad with VR glasses, 5G mobile phone and a PC with 5G modem. There can naturally be much more receiving clients as NGINX is capable to deliver hundreds of parallel streams.

Sending and receiving delays for the streams are to be measured. Several parallel streams are not needed to be able to measure the delay on the sending or on the receiving side. Therefore, only one uplink stream can be used. Same goes for the receiving side and also there, only one downlink stream is enough. Figure 26 illustrates the generic test setup.

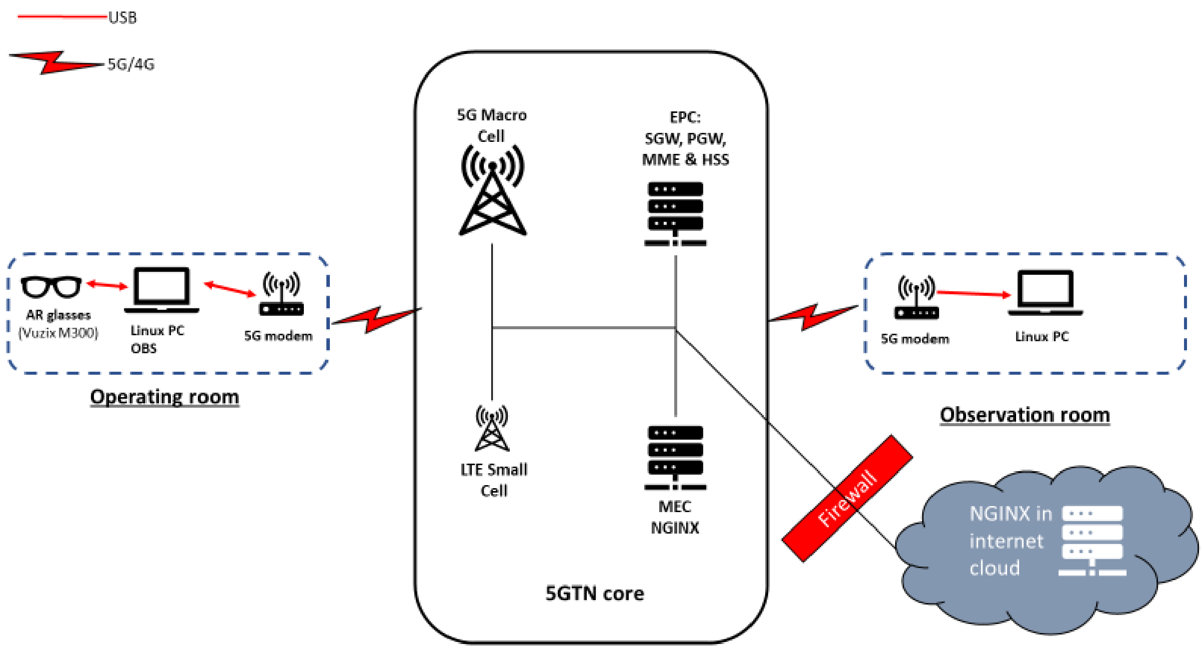


Figure 26. Generic test setup.

Scope of this thesis is to measure data traffic delays to and from the NGINX multimedia server. All of the equipment and the streams that are used in the hospital demo are not needed to reach that goal. Therefore, the system is simplified but the essence of the use case is still maintained. There is only one stream that is being sent to the multimedia server using 4G/5G connection and there is one computer receiving the stream through 4G/5G connection. There are two possible locations for the NGINX multimedia server as described in Figure 26: Edge server or internet cloud. In the sending side camera source can be AR glasses or another camera that feeds continuous data at same data rate as the AR glasses.

4.3 Main test setup and components

All the test setup components and their connections are explained in this chapter. Both HW and SW that is used is explained.

4.3.1 Operating room

Following equipment is used in the operating room:

AR glasses [41]:

- Vuzix M300 XL AR glasses
- Android 6.0 Operating System (OS)
    - Vuzix dependent
    - No upgrade possibility
- 10 Mega Pixel Camera
- 1080p video
- USB Micro-B 2.0
- WiFi b/g/n/ac – Dual-Band 2.4/5 GHz
- Hospital Demo App by University of Oulu

Camera

- Can be used instead of the AR glasses
- Razer Kiyo broadcasting camera with illumination
- USB 2.0 connection
- Up to 1080p at 30 frames per second (fps) video capability
- Can be used instead of the AR glasses

PC

- Dell Latitude E6420 Laptop PC
    - Ubuntu Linux version 18.04 Long Term Support (LTS).
- Open Broadcaster Software (OBS) version 24.0.3 for Windows

5G modem

- Mediatek 5G modem
- Supports 4G and 5G connectivity that are needed in this thesis

Operating room equipment is connected so that AR glasses are connected to the PC using USB 2.0 cable. Mediatek 5G modem is connected to the PC using USB 3.0 cable. Mediatek modem connects to the 5GTN using 4G or 5G connection depending on the test case.

4.3.2 5G Test Network

5G Test Network works in a Non-stand-alone mode. See chapter 2.5.1 5G Test Network for more details about NSA. It has several Nokia Flexi Zone Multiband Indoor Pico BTS 4G base stations and two Nokia 5G base stations operating at 3.5 GHz frequency. Core network has also normal EPC functionality (PGW, SGW, MME and HSS) as well as several Edge Servers. 5GTN is also connected to internet through PanOulu network.

NGINX was installed on one of the Edge Servers (Nokia Airframe with Intel Xeon E5-2630 processor) that was running Ubuntu Linux version 18.04 LTS. NGINX version used was 1.16.1 together with RTMP module.

4.3.3 Observation Room

Scenario here is that the video stream is received through 5G modem to the PC where it can be played by a compatible video player like VideoLAN client (VLC). Stream could also be viewed from a big screen or a large Television (TV). Another option is to receive the stream directly with 5G mobile phone and view the stream in the phone or for example through a TV connected to the phone.

Observation room equipment is simplified for testing purposes to include only one PC receiving the stream from the NGINX multimedia server. Following equipment was used:

- Dell Latitude E4310 laptop
  - Linux 18.04 LTS operating system
- VLC media Player
- Mediatek 5G modem.

Stream was also received using Huawei Mate 20 X 5G mobile phone.

4.3.4 Cloud Service

NGINX was installed to Edge Server in the 5G network as baseline. Reference installation was made to Microsoft Azure. This reference implementation presents practically the only way to locate the NGINX multimedia server before it can be installed to Edge Server. These two systems with different NGINX locations are then compared to each other. Azure environment was like this:

- Size
  - Standard B1ms (1 virtual central processing unit (CPU), 2GB memory)
- Operating system
  - Ubuntu Linux 18.04 in virtual environment
- Location
  - North Europe
- Networking with external ip
- Precision Time Protocol daemon (PTPd) installed
- NGINX installed
- Linux access through remote terminal

4.4 Network measurement equipment

To be able to measure transmission delays in millisecond level sets high requirements for the test system and the equipment. One must be able to measure extremely short delays on a reliable way so that measurement results between each test are comparable and repeatable.

Kaitotek's Qosium software was used to measure the transmission delays in the test system. Qosium is a passive Quality of Service (QoS) and Quality of Experience (QoE) performance measurement and monitoring software. Qosium does passive measurements in real time and also presents the measurements results in real time. [42]

Qosium measurement system is built so that there is a Qosium Probe installed on network devices on both ends of the system where the end-to-end measurements are to be done. Qosium Scope analyser SW is then used for creating, controlling, and parameterizing Qosium measurements. Qosium Scope also collects and visualizes results. [42]

In test system where the NGINX is located at the 5GTN Edge Server, Qosium probes are installed to the following locations:

- on the Operating room PC that sends the stream to the NGINX
- to the Edge Server where the NGINX is installed
- to the Observation room PC that is receiving the stream from the NGINX.

One set of measurements can be done between two probes only. In this case between probes in Operation room PC and Edge server or between Edge server and Observation room.

This kind of system is described in Figure 27 below. Data transmission time between Operating and Observation rooms can’t be measured directly as the data packets are not sent between those two locations directly but NGINX in the Edge server processes the stream before it is sent as a new stream to Observation room.

Measuring very short millisecond level transmission delays between two separate computers sets also high requirements for clock synchronization. Windows OS’s clock resolution and accuracy are not on a level that suits the needs of these measurements. There is not only a low resolution, but the clock also drifts to some direction continuously causing increasing error to the measurements. Therefore, Linux - that has better clock accuracy than Windows – was needed to be used in all environments where Qosium probes were installed. Linux clock service was improved by installing Precision Time Protocol (PTP) service to all Linux machines. PTP clock resolution is on microsecond level and that level of accuracy improves the reliability of these measurement results. [43]

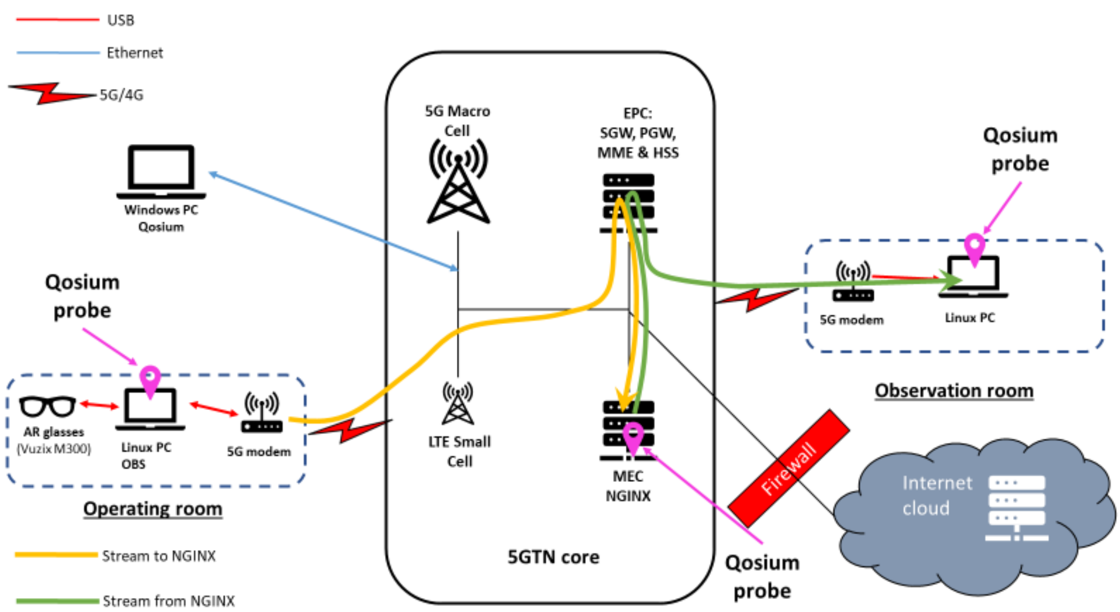


Figure 27. NGINX in Edge Server test setup.

When NGINX is installed on the cloud server instead of the Edge Server the Qosium probe installations are the same in both systems except that now probe is installed to the Cloud server instead of the Edge Server. Also here the tests are performed for separately for sending the stream (between Operating room and Edge server) and for receiving the stream (between Edge server and Observation room. Cloud Server test system is described in Figure 28.

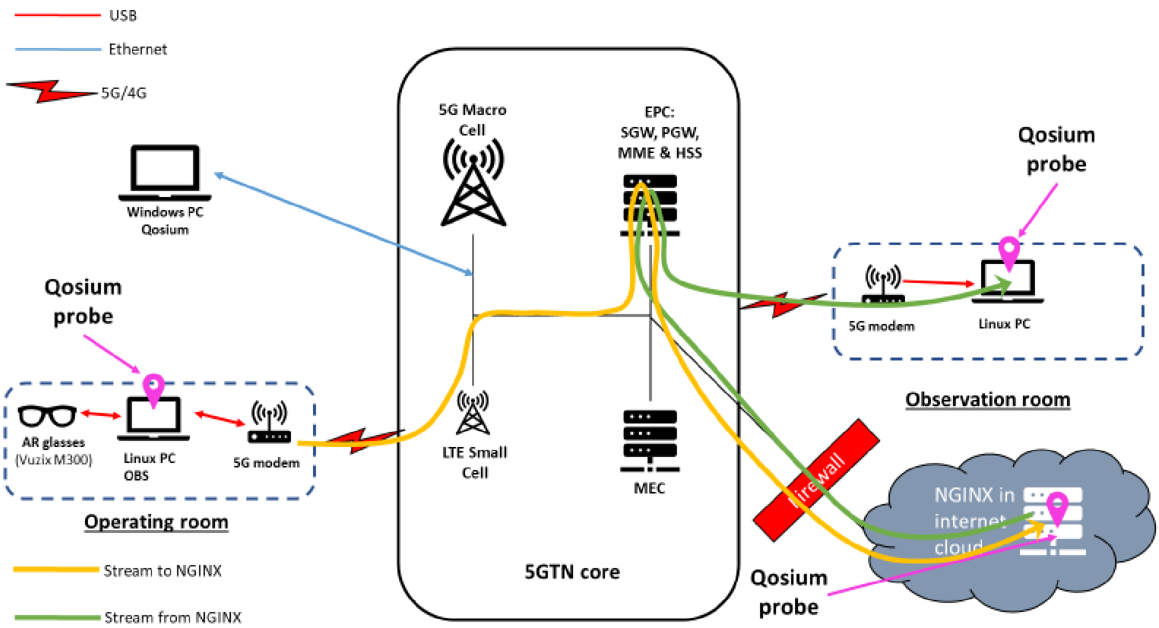


Figure 28. NGINX in Cloud Server test setup.

5 TESTING PROCEDURE AND RESULTS

5.1 Test specification

Test case specification for the tests to be done was created. Purpose of the test document is to secure the repeatability and comparability of tests if they need to be done several times and also between different configurations of one test session.

Following steps were defined (see Appendix 1 for further details):

- NGINX at Edge Server. 5G.
- NGINX at Edge Server. 4G.
- NGINX at Cloud Server. 5G.
- NGINX at Cloud Server. 4G.

Basic knowledge of the system in question is assumed to be known from the person executing the tests. Test case specification is to be followed in all the tests. If specification can't be followed, the tests fail for a known reason or there is need then for the specification to be updated. If the specification is updated all the tests done so far are to be done again if the specification change could have an impact on the test results.

5.2 Testing operations

Edge Server related tests were executed at the University of Oulu. Testing system presented in Figure 27 was built:

- Sending side system in the Operating room included laptop PC where Ubuntu Linux was installed. Mediatek 5G modem including the sw drivers was installed to the laptop. OBS SW and PTP clock synchronization SW were also installed. HP Webcam HD 2300 was connected to the laptop.
- NGINX media server SW was installed to Edge Server and the system was configured to receive RTMP stream and to send a stream both in RTMP and HLS formats. RTMP protocol was used in the testing for both send and receive. PTP clock synchronization SW was installed to Edge Server.
- Receiving side system in the Observation room included laptop PC where Ubuntu Linux was installed. Mediatek 5G modem including the SW drivers was installed to the laptop. VLC video playback SW and PTP clock synchronization SW were also installed.
- PTP clock synchronization SW was configured to be used in the unicast mode so that the link that was measured was active. As the stream delay was measured the uncast sync link was active between the laptop in the Operating room and the Edge server. When the stream receiving delay was measured, the synchronization was done between the Edge Server and the receiving laptop in the Observation room. Edge server was always the PTP clock master and the laptops were in slave mode.
- Qosium measurement PC had Windows 10 installed. Qosium Scope SW was installed, and the PC was connected to 5GTN using ethernet cable.
- Qosium probes were installed as service to each Linux machine in the system: sending laptop, receiving laptop and the Edge Server.

Tests were executed according to test specification in the following way:

- A location where there was a line of sight between the sending and receiving PCs' 5G modems and the 5G antenna was used.
- Video stream sending and receiving laptops were powered on.
- 5G connection for both laptops was established using the 5G modems.
- PTP clock sync was started between sending side laptop and the Edge Server.
- NGINX multimedia server was started at the Edge Server.
- HP webcam video stream was sent to multimedia server using OBS SW. Video bandwidth was limited to 600 kilobits per second (kbit/s).
- Video stream was received using VLC SW in the receiving laptop.
- Qosium Scope in the measurement PC was started.

5G Measurements for sending the video:
- Qosium Scope was connected to sending laptop and Edge Server Qosium Probes.
- Preliminary measurement was started to see that clocks are synchronized between the sending laptop and the Edge Server.
- When clock synchronization was achieved the actual measurement was started.
- The first 5 min measurement interval was executed.
- Measurement results of the 5 min test were saved to measurement PC.
- A second measurement, a 1-minute interval was started.
- Measurement results of the 1 min test were saved to measurement PC.

5G Measurements for receiving the video:
- Qosium Scope was now connected to probes in the Edge Server and in the receiving laptop.
- Clock synchronization was established between the Edge Server and the receiving laptop.
- Preliminary measurement was started to see that clocks are synchronized between the sending laptop and the Edge Server.
- When clock synchronization was achieved the actual measurement was started.
- The first 5 min measurement interval was executed.
- Measurement results of the 5 min test were saved to measurement PC.
- A second measurement, a 1-minute interval was started.
- Measurement results of the 1 min test were saved to measurement PC.

Measurements for 4G connection:
- A location where 5G signal was weak needed to be found. When 5G signal is weak the modem uses 4G connection instead of the 5G. Tests were executed in that location using only 4G.
- Tests themselves are same as for the 5G but now the air interface is 4G instead of the 5G.
- Test specification was followed.
- Test results were stored to measurement PC.

Measurements when NGINX multimedia server is in the internet Cloud:
- Same as for Edge Server but now the NGINX is in the internet cloud server instead of the Edge Server.

- Tests were executed for both 4G and 5G connections using the same locations as for the Edge Server measurements.
- Test specification was followed with two exceptions.
    - When measuring sending of the video:
        - Sending side PC was connected through 4G or 5G connection to 5GTN.
            - This connection was under measurement
        - The PC receiving the video used ethernet connection to 5GTN.
            - This connection was not measured
    - When measuring receiving the video
        - Receiving side PC was connected through 4G or 5G connection to 5GTN.
            - This connection was under measurement
        - The PC sending the video used ethernet connection to 5GTN
            - This connection was not measured
- Test results were stored to measurement PC.

5.3 Test results

All the measurement results are presented here. Delay is shown as a graph presenting the delay as a function of time. Delay is shown over the whole measurement period which (as set in the Test Specification) is 1 minute or 5 minutes. Graph for both durations, 1 minute and 5 minutes, is presented.

In these measurements Qosium measurement tool shows average results with one second resolution. This means that Qosium measures the data for one second, then averages the results of it and presents that as a measured value. Then same procedure is repeated for the next one second long period and so on. These results are then presented in graphical form. Those graphs are presented here.

Qosium measurement results were also stored in text format so that that the results can be presented and analysed also after the measurements. These text format files were imported to Excel where also an average delay over the whole 1 minute or 5 minutes measurement period was calculated. That average delay is below presented alongside the graphs.

Every graph includes bandwidth measurement for uplink and downlink channels on the left side and delay measurement information on the right side. Note the direction of the video. When sending the video, it is occupying the uplink channel and downlink is used only for control messages. When receiving the video, it is on the downlink channel and the control signalling is on the uplink channel.

5.3.1 4G Edge server measurements

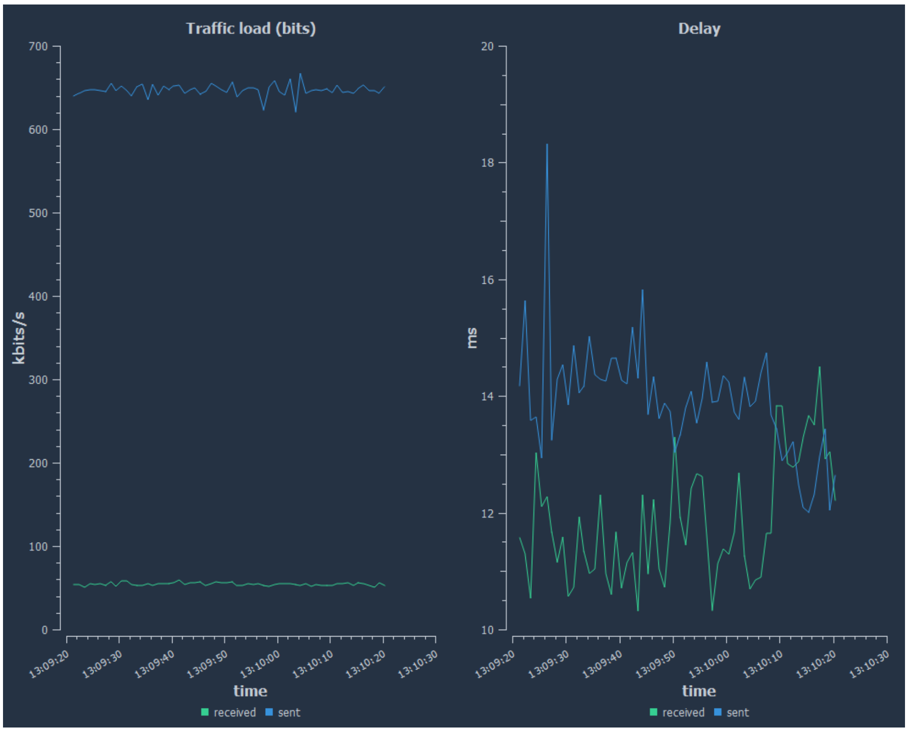


Figure 29. 4G connection, Edge server, video receive, 1-minute interval.

Results of the Figure 29 average delays: uplink 11.84 ms, downlink 13.92 ms

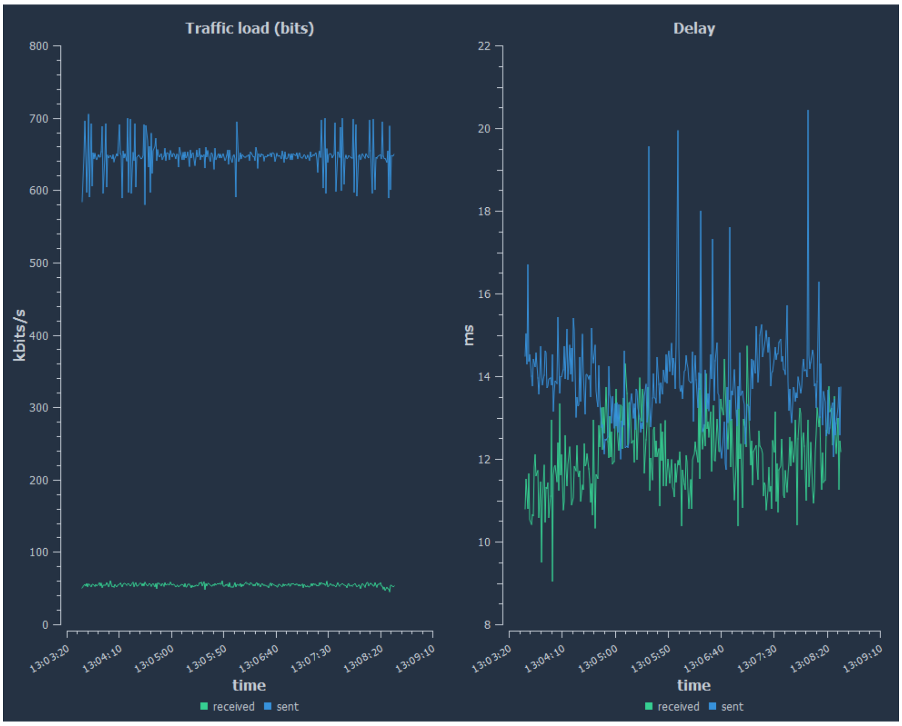


Figure 30. 4G connection, Edge server, video receive, 5-minute interval.

Results of the Figure 30 average delays: uplink 12.09 ms, downlink 13.86 ms

Figure 31. 4G connection, Edge server, video send, 1-minute interval.

Results of the Figure 31 average delays: uplink 22.71 ms, downlink 13.68 ms

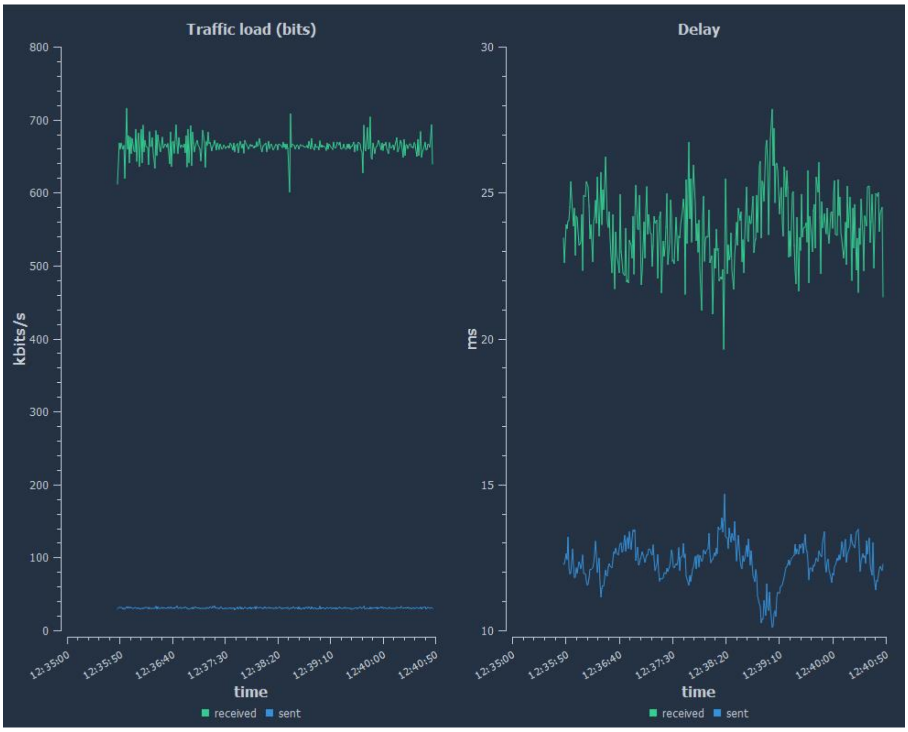


Figure 32. 4G connection, Edge server, video send, 5-minute interval.

Results of the Figure 32 average delays: uplink 23.83 ms, downlink 12.34 ms

5.3.2 5G Edge server measurements

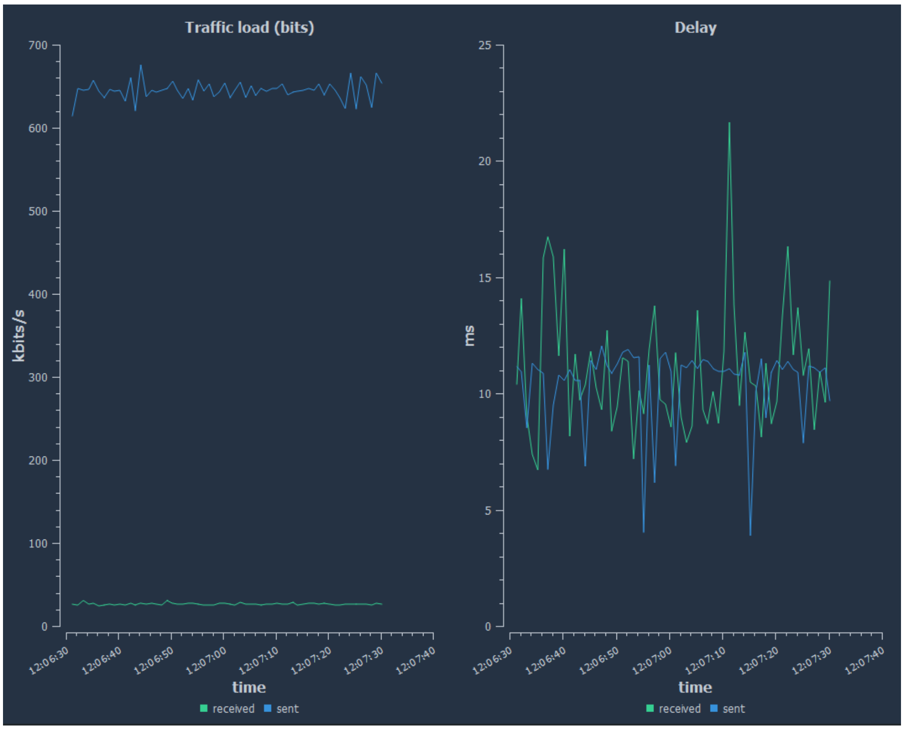


Figure 33. 5G connection, Edge server, video receive, 1-minute interval.

Results of the Figure 33 average delays: uplink 11.10 ms, downlink 10.43 ms

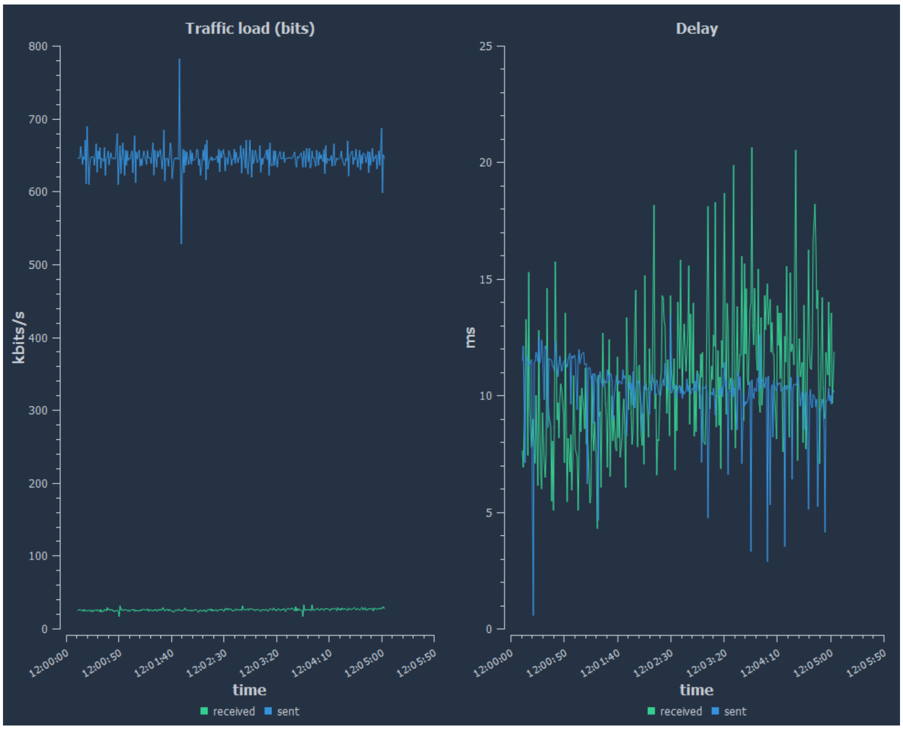


Figure 34. 5G connection, Edge server, video receive, 5-minute interval.

Results of the Figure 34 average delays: uplink 9.90 ms, downlink 10.40 ms

Figure 35. 5G connection, Edge server, video send, 1-minute interval.

Results of the Figure 35 average delays: uplink 11.35 ms, downlink 12.89 ms

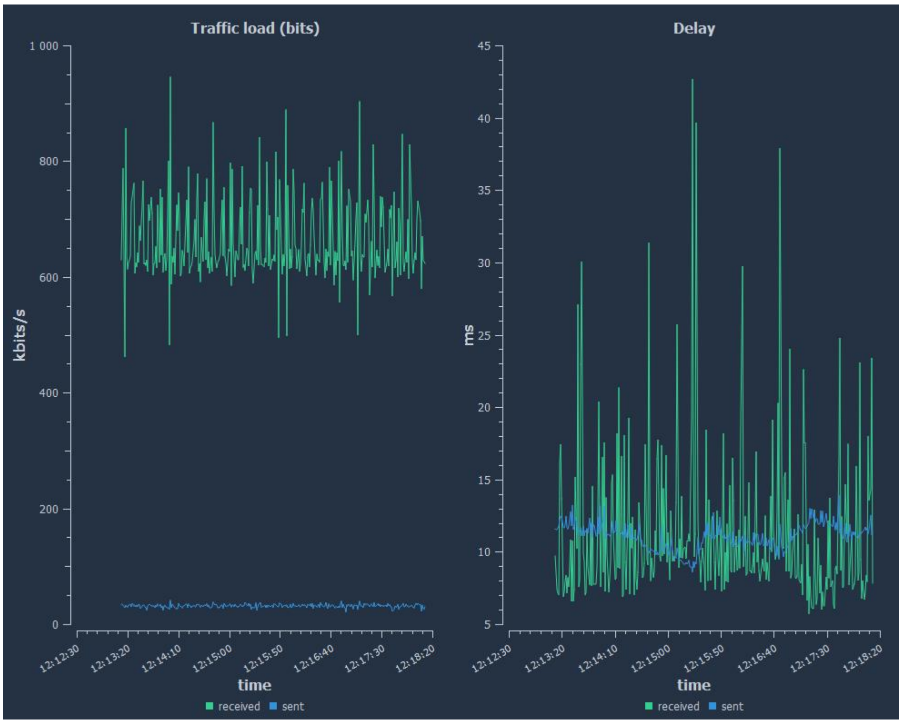


Figure 36. 5G connection, Edge server, video send, 5-minute interval.

Results of the Figure 36 average delays: uplink 11.06 ms, downlink 11.26 ms

5.3.3 4G Cloud server measurements

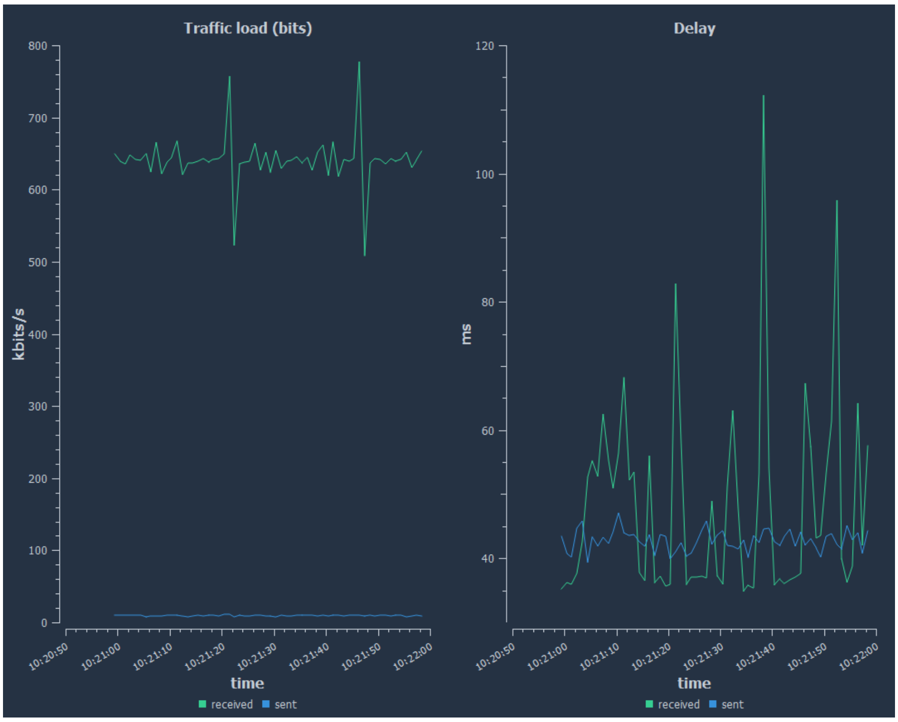


Figure 37. 4G connection, Cloud server, video receive, 1-minute interval.

Results of the Figure 37 average delays: uplink 47.96 ms, downlink 42.78 ms

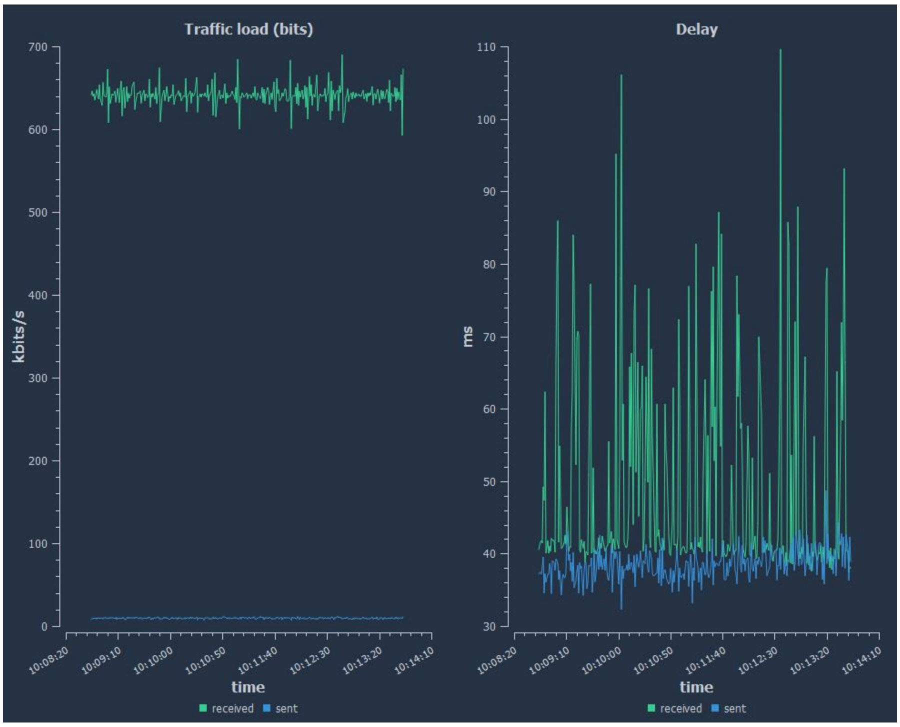


Figure 38. 4G connection, Cloud server, video receive, 5-minute interval.

Results of the Figure 38 average delays: uplink 38.54 ms, downlink 48.98 ms

Figure 39. 4G connection, Cloud server, video send, 1-minute interval.

Results of the Figure 39 average delays: uplink 54.27 ms, downlink 35.94 ms

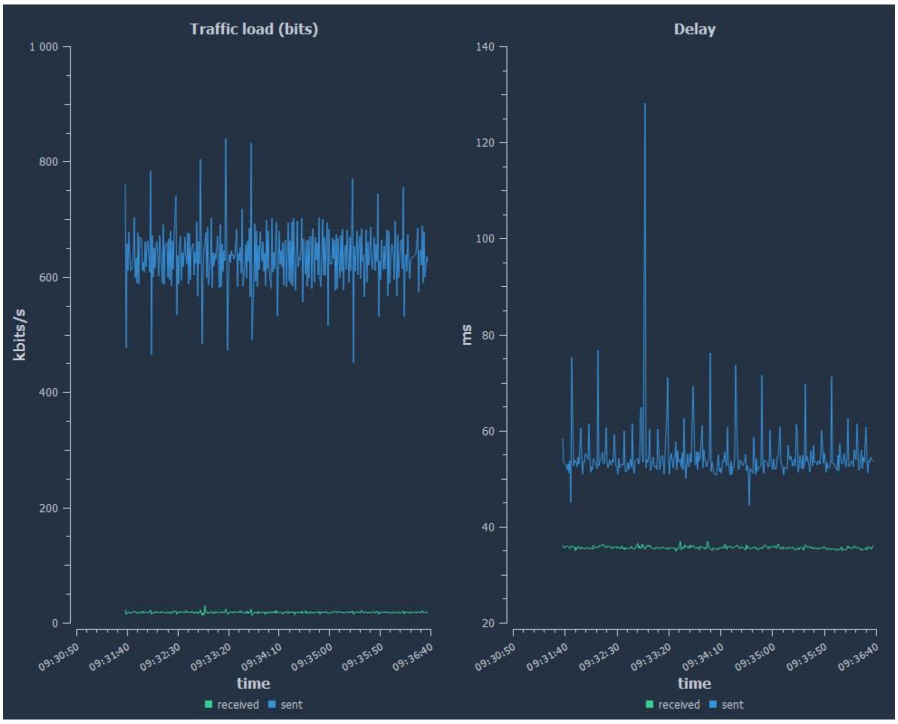


Figure 40. 4G connection, Cloud server, video send, 5-minute interval.

Results of the Figure 40 average delays: uplink 54.63 ms, downlink 35.60 ms

5.3.4 5G Cloud server measurements

Figure 41. 5G connection, Cloud server, video receive, 1-minute interval.

Results of the Figure 41 average delays: uplink 31.40 ms, downlink 44.41 ms

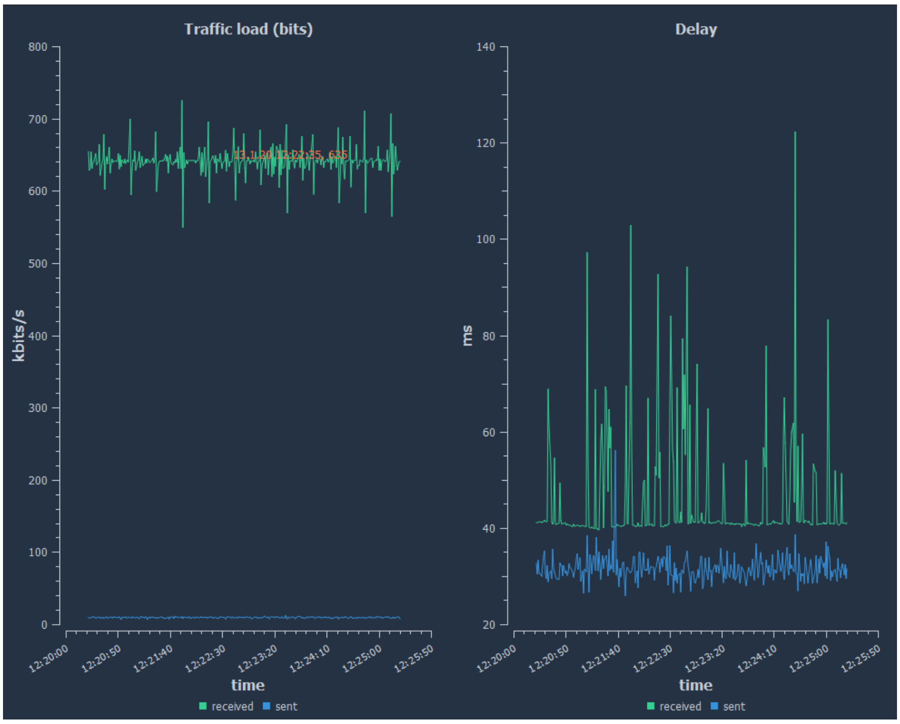


Figure 42. 5G connection, Cloud server, video receive, 5-minute interval.

Results of the Figure 42 average delays: uplink 31.55 ms, downlink 45.74 ms

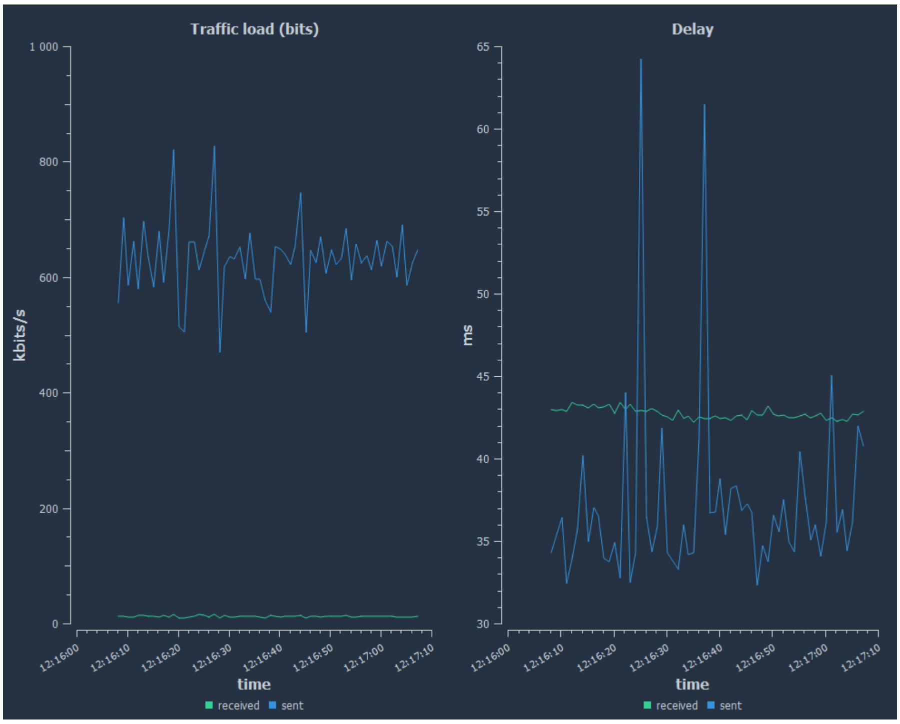


Figure 43. 5G connection, Cloud server, video send, 1-minute interval.

Results of the Figure 43 average delays: uplink 37.15 ms, downlink 42.73 ms

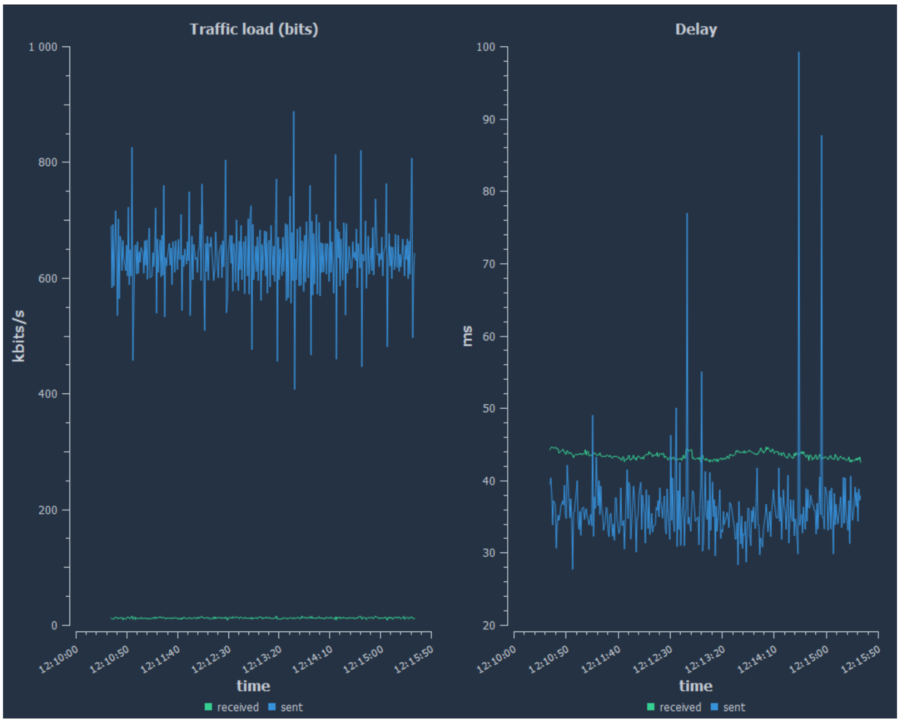


Figure 44. 5G connection, Cloud server, video send, 5-minute interval.

Results of the Figure 44 average delays: uplink 36.05 ms, downlink 43.40 ms

5.3.5 Test System Stability

To test the stability of the testing environment, the test system was kept actively sending and receiving the stream until it was interrupted. Continuous stream went on for 217 hours, 30 minutes and 33 seconds until modem connection was dropped.

During the testing the test system was very stable, and no crashes happened during the measurements. The only stability issue that was observer was the difficulty to create a stable connection between the Linux laptops and the 5G modems. As the Linux driver software was quite early version the driver software needed to be run several times before a stable permanent connection was established. When the connection was finally established it then stayed on without any problems.

6 ANALYSIS OF TEST RESULTS

In this chapter the delay measurements are analysed. First the measurements done using the 4G and the 5G connections are compared to each other. Purpose is to understand the differences these two systems have. Next, the systems that differ in the location of the NGINX installation are compared. Target is to analyse the impact the Edge server usage has on the delay when comparing it with internet cloud alternative. Lastly the conclusion based on the measurements are done.

6.1 Comparison 4G vs. 5G

There were four different measurements done using the test system with 4G connection when NGINX was installed on the Edge server. Similarly, four measurements were done with 5G connection and NGINX on the Edge server.

6.1.1 Edge server

The 1- and 5-minute tests do show practically the same delays, so there isn't much deviation between different tests with different durations. In Receive cases below, video is streamed form the NGINX multimedia server to the Observation Room PC. In send cases video is being sent from the Operating Room PC to the NGINX multimedia server.

Measurements with 4G connection has the following delay values

- Receive:
  - 1 minute: uplink 11.84 ms and downlink 13.92 ms
  - 5 minutes: uplink 12.09 ms and downlink 13.86 ms
  - Uplink difference: 0.25 ms
  - Downlink difference: 0.06 ms
- Send:
  - 1 minute: uplink 22.71 ms and downlink 13.68 ms
  - 5 minutes: uplink 23.83 ms and downlink 12.34 ms
  - Uplink difference: 1.12 ms
  - Downlink difference: 1.34 ms

Measurements with 5G connection has the following delay values

- Receive:
  - 1 minute: uplink 11.10 ms and downlink 10.43 ms
  - 5 minutes: uplink 9.90 ms and downlink 10.40 ms
  - Uplink difference: 1.20 ms
  - Downlink difference: 0.03 ms
- Send:
  - 1 minute: uplink 11.35 ms and downlink 12.89 ms
  - 5 minutes: uplink 11.06 ms and downlink 11.26 ms
  - Uplink difference: 1.12 ms
  - Downlink difference: 1.34 ms

All average delay for each option is gathered to Table 3 below:

Table 3. Comparison of the measured delays (Edge server)

			5G difference to 4G (ms)	
	Send (ms)	Receive (ms)	Send (ms)	Receive (ms)
4G uplink 1 min	22.71	11.84	-	
4G downlink 1 min	13.68	13.92	-	
4G uplink 5 min	23.83	12.09	-	
4G downlink 5 min	12.34	13.86	-	
5G uplink 1 min	11.35	11.10	-11.36	- 0.74
5G downlink 1 min	12.89	10.43	-2.33	-3.49
5G uplink 5 min	11.06	9.90	-12.77	-2.19
5G downlink 5 min	11.26	10.40	-1.08	-3.46

There is a surprisingly small difference between 4G and 5G downlink delays when receiving the video stream: an average of 3.49 milliseconds between 5-minute measurements. One might assume the difference to be bigger but as the video bandwidth is limited to be only just above 600 kb/s, both the 4G and 5G downlink channels can easily transfer the video to the client through the downlink channel.

Uplink has the difference in delay for 2.19 ms in the same 5-minute video receiving measurement. There is only a small amount of control data (a few tens on kb/s) transferred in the uplink channel from the streaming client to the multimedia server. These packets can be transferred fast in the uplink channel even though it has smaller bandwidth than downlink in the 5G Test Network.

Larger difference between the 4G and the 5G connections is visible when sending the video stream uplink to the Edge server. 4G delay is more than double comparing to 5G delay: from 11.06 ms to 23.83 ms in the 5-minute measurement. This is most probably due to the uplink channel being much narrower than the downlink channel. During these tests the bandwidth ratio between uplink and downlink was one to seven, uplink being the narrower. It seems obvious that the video has longer uplink transfer time in 4G than in 5G due to the 5G having much better uplink bandwidth capacity.

6.1.2 Cloud server

The 1- and 5-minute tests do have a noticeable difference when receiving the video using 4G connection, so the 5-minute test results are used for the analysis as they are seen to be more reliable due to the longer test duration. The measurement results in the sending side are very close to each other.

Measurements with 4G connection has the following delay values

- Receive:
    - 1 minute: uplink 47.96 ms and downlink 42.78 ms
    - 5 minutes: uplink 38.54 ms and downlink 48.98 ms
    - Uplink difference: 9.42 ms
    - Downlink difference: 6.2 ms
- Send:
    - 1 minute: uplink 54.27 ms and downlink 35.94 ms

- 5 minutes: uplink 54.63 ms and downlink 35.60 ms
- Uplink difference: 0.36 ms
- Downlink difference: 0.34 ms

Measurements with 5G connection has the following delay values

- Receive:
  - 1 minute: uplink 31.40 ms and downlink 44.41 ms
  - 5 minutes: uplink 31.55 ms and downlink 45.74 ms
  - Uplink difference: 0.15 ms
  - Downlink difference: 1.34 ms
- Send:
  - 1 minute: uplink 37.15 ms and downlink 42.73 ms
  - 5 minutes: uplink 36.05 ms and downlink 43.40 ms
  - Uplink difference: 1.10 ms
  - Downlink difference: 0.67 ms

All the average delay for each option is gathered to table 3 below:

Table 4. Comparison of the measured delays (cloud server)

			5G difference to 4G (ms)	
	Send (ms)	Receive (ms)	Send (ms)	Receive (ms)
4G uplink 1 min	54.27	47.96	-	
4G downlink 1 min	35.94	42.78	-	
4G uplink 5 min	54.63	38.54	-	
4G downlink 5 min	35.60	48.98	-	
5G uplink 1 min	37.15	31.40	-17.12	- 16.56
5G downlink 1 min	42.73	44.41	+6.79	+1.63
5G uplink 5 min	36.05	31.55	-12.77	-6.99
5G downlink 5 min	43.40	45.74	+7.80	-3.24

Here the measurements do show that 5G connection would be slower than the 4G in 3 out of 4 comparisons in the downlink direction. That is a clear change from the results of the similar measurements when the NGINX was in the Edge server. The reason for the difference can’t be analysed from the measurement results themselves. The reason could be in the load difference of opposite data flow directions of internet or Azure. On the other hand, the measurements are done in a different time: 4G from around 9.30 AM until 10.20 AM and 5G from 12.10 until 12.27. There could also be higher internet traffic towards the Oulu university network from the internet as the 5G measurements were done during the lunch hours and 4G before that.

It is also possible that there is a difference in the clock synchronization between the measurements. Still it can be fairly safely assumed that the difference of the results is not due to the 5G itself, but other reasons or difference would have been visible in the Edge server measurement results presented in the chapter 6.1.1 Edge Server. Focus of this thesis is to compare Edge server to cloud server, not 4G to 5G. Comparing the 4G to the 5G is an additional outcome that was got from the measurements. Therefore, the reasons behind the differences are not investigated further.

6.2 Comparisons of Edge Server vs. cloud

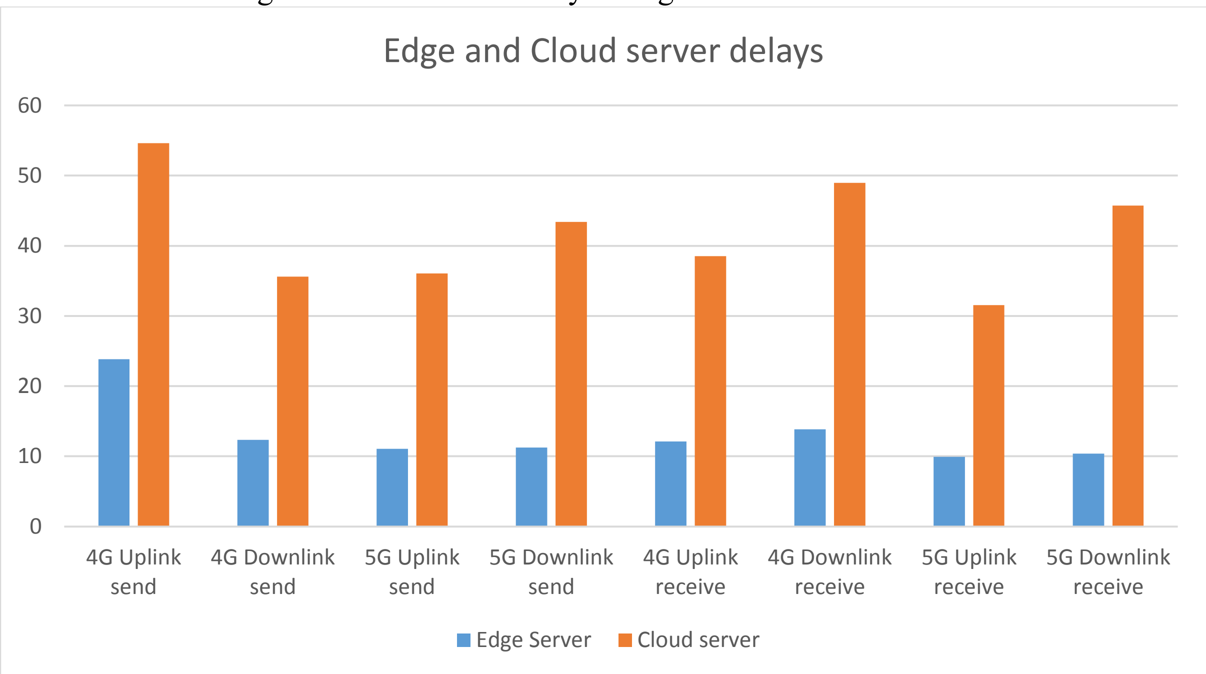
Here the measurements done for the option where NGINX was installed on the cloud server are compared against the option where NGINX is installed on the Edge server close to the base station. Four different measurements were done using the test system with 4G connection: 1-minute and 5-minutes measurements when the NGINX is in the cloud server and the same for the option when NGINX is in the Edge server. Same four measurements were done for 5G connection also.

Edge server and cloud server related delays and their differences are presented in Table 5 and Table 6 below. Figures are from 5-minute measurements.

Table 5. Edge and cloud server delays using 4G and 5G connections

	Send		Difference		Receive		Difference	
	Edge server	Cloud server	ms	%	Edge server	Cloud server	ms	%
4G uplink	23.83	54.63	30.80	129	12.09	38.54	26.45	219
4G downlink	12.34	35.60	23.26	188	13.86	48.98	35.12	253
5G uplink	11.06	36.05	24.99	226	9.90	31.55	21.65	219
5G downlink	11.26	43.40	32,14	285	10.40	45.74	35.34	340

Table 6. Chart of edge and cloud server delays using 4G and 5G connections



As it can be seen from Table 5 above there is a clear difference in the delays between Edge and cloud server options. Delay is from 21.65 ms (5G uplink when receiving the video) up to 35.34 ms (5G downlink when receiving the video) longer when using the Cloud server than when using the Edge server. Percentage wise the difference is from 129% (4G uplink sending the video) all the way up to 340% (5G downlink receiving the video).

6.3 Conclusions

With 5G connection the transmission delay for Edge server was as low as 9.90 milliseconds and even with 4G as low as 12.09. Comparing to 31.55 ms and 38.54 ms with Cloud server using same test cases. Cloud server adds from 21.65 ms to 35.34 ms to the transmission delay of the video stream depending whether it is utilizing 4G or 5G connection and what is the direction of the video stream. There is also a clear difference between 4G and 5G connections especially in the 5-minute uplink video stream case where 4G has a measured average delay of 23.83 ms and 5G a measured average delay is only 11.06 ms.

It can be concluded, based on executed delay measurement results, that there is a clear benefit of utilizing the Edge server for multimedia server location instead of the Cloud server. If the benefits are good enough to install the application to the Edge server instead of the Cloud server depends on the operators, enterprises or persons responsible for respective functionality.

If there is a need for fast response time and/or minimum delay when using 4G or 5G connectivity, Edge server can help delivering that. In fact, Edge server in 5G network can offer the shortest data transmission delays between the client and server from all the current mobile phone networks.

7 DISCUSSION

This thesis was a continuation of the Bachelors work done during the late 2018 and the first half of the 2019. In the Bachelors work Hospital Use Case Demo the first version was defined and created. In the first demo, a video stream was sent through (Wireless Fidelity) WiFi connection to NGINX and then streamed to an iPad where it was shared to VR glasses. 5G connection was delivered through 5G WiFi hot spot.

The first demo developed was then used as a basis for this thesis. Demo was further developed to suit the needs of the demo requirements and this thesis. Most importantly the hospital Use Case Demo system needed to be made 5G capable. First the NGINX multimedia server was installed on an 5GTN Edge server and configured to work with the demo that was still using the WiFi connection. Then demo was made 5G capable by integrating Mediatek 5G modems to the system. Usage of separate cameras needed to be defined as the AR glasses and 360° cameras were not very practical for the testing purposes.

A new test system needed to be defined and taken into use. Qosium Test system from Kaitotek was found to fulfil the testing needs. First demo was working on top of Windows Operating System, but Windows' clock resolution is not good enough for Qosium to perform millisecond class delay measurements but the one in Linux is. Therefore, the whole demo system needed to be transferred to work on top of Linux operating system. Linux clock resolution was further enhanced by installing separate PTP clock synchronization software to all Linux machines.

Initial intention was to utilize MEC application installed to Edge server as the location for NGINX in the 5GTN. That concept was developed as a part of a project in the University of Oulu and the purpose was to offer the MEC as platform for the applications. As the work progressed it became clear that MEC platform would not be available in the time frame of this thesis. So, a decision was made to install the NGINX as pure Linux application directly to the Edge server.

NGINX needed to be installed also to a cloud server. Microsoft Azure was selected to be used as the cloud service as it is a reliable large operator in the cloud service area. Linux, NGINX as well as PTP clock sync SW and Qosium probe were installed to Azure.

A lot of the different technologies, devices and software were new not only to the author but to 5GTN personnel also. No one had used them before, and they needed to be made operational as part of this thesis. For example, 5G modem was first time made to connect to 5GTN and Qosium measurement system was built. The amount of new technology proved to be the most challenging part of the work done for the thesis as a lot of time needed to be spent to learn and make the technology functional.

Results of the tests were as expected: Edge server delays are smaller than Cloud server delays. That is natural as the physical distance between the base station and Edge server is very small compared to distance to cloud server. From that perspective the goals for the thesis were achieved.

It was also shown that 5G delays are smaller than 4G delays, so the 5G can be said to fulfil its promise. Further improvement in 5GTN performance can be expected when moving from the Non-Stand-Alone system to purely Stand-Alone 5G system with own 5G core components.

8 SUMMARY

Both the amount of data transferred and the number of connected devices of the mobile phone networks is multiplying during the coming 4 years as shown in Figures 1 and 2. New mobile phone network generation 5G is being deployed to the market from 2019 onwards. It will enable more devices to be connected to the network as well as introduce new technologies, higher data rates and shorter delays. One of the new technologies 5G brings is Edge computing that brings cloud computing type of computing inside the 5G network close to the base station. Edge server together with low latency communication enables fast communication between user equipment and the server application. This thesis focuses on measuring the delay difference between the Edge and cloud server technologies.

5G network is introduced in 3GPP releases from Release 15 onwards. Releases 16 and 17 are now planned to include 5G features. 5G architecture will introduce drastic changes to the mobile phone network architecture through its main design principles that introduce slicing, SDN and NFV. By those design principles formerly HW based modules are transformed to be SW based within the 5G network. This enables new functionality and new business models and opportunities not only for the operators but also for SW platform and application creators.

This thesis uses Hospital Use Case demo that has been developed in the University of Oulu as basis for the test environment. This demo is built for remote consultation, home hospital and education purposes. Hospital Use Case demo is further developed to include the features and functionalities need to perform the required measurements. Demo was, for example, modified to use 5G connections for the video streaming and the computing environment was changed from Windows to Linux.

The actual test environment included a camera, Linux laptop and a 5G modem in the operating room sending the video stream. Multimedia server receiving the stream resided in an Edge server in the 5G Test Network. Laptop connected to 5G Test Network received the stream from 5GTN and acted as Operating room equipment.

Tests were done using Qosium measurement system and they were executed according to the test specification that was created for the measurements. Live video stream was sent to multimedia server at the edge server and then streamed form the multimedia server. Delay measurements were done using both 4G and 5G connections and for downlink and uplink directions when sending and receiving the video. Same tests were done for 1-minute and for 5-minute durations. Then multimedia server was installed to internet cloud and the measurements were repeated. All measurements results were stored for further analysis.

Results were compared and the edge server delays were found to be much shorter than cloud server delays. With 5G, the average delay was about 10 milliseconds and with cloud server, approximately 31 milliseconds. When using 4G connection, the delays were around 12 and 39 milliseconds, respectively. The measurement results conclude, that here is a clear benefit utilizing the Edge server for multimedia server location instead of the Cloud server, irrespective of the used mobile connection, 4G or 5G.

9 REFERENCES

[1] ITU-T L Suppl. 36, ITU-T L.1310 – Study on methods and metrics to evaluate energy efficiency for future 5G systems, supplement 36, (11/2017) (2018), Telecommunication Standardization sector of ITU, Geneva, Switzerland, 24 p.

[2] Warren D. & Calum D. (2014), Understanding 5G: Perspectives on future technological advancements in mobile, GSMA Intelligence, The Walbrook Building, 25 Walbrook, London, https://www.gsmaintelligence.com/research/?file=141208-5g.pdf&download, 26 p.

[3] Future mobile data usage and traffic growth (read 29.7.2019), Ericsson, https://www.ericsson.com/en/mobility-report/future-mobile-data-usage-and-traffic-growth

[4] Ericsson Mobility Visualizer (read 29.7.2019), https://www.ericsson.com/en/mobility-report/mobility-visualizer?f=7&ft=2&r=1&t=1,20&s=4&u=3&y=2014,2024&c=3

[5] The Promise and potential of 5G: Evolution or revolution (2019) (read 31.7.2019), IHS Markit, https://cdn.ihs.com/www/pdf/0419/IHS-Markit-Technology-5G-The-Promise-Potential.pdf, 13 p.

[6] Edge cloud web pages (read 20.1.2020), Nokia , https://www.nokia.com/networks/solutions/edge-cloud/

[7] 3GPP web pages (read 1.8.2019), 3GPP, https://www.3gpp.org/about-3gpp

[8] 3GPP releases web page (read 1.8.2019), https://www.3gpp.org/specifications/67-releases

[9] Bertenay B. (2019) 5G Standardization update, https://www.3gpp.org/ftp/Information/presentations/presentations_2019/2019_07_Brighttalk_Post-RAN84.pdf, 3GPP, 14p.

[10] 5G Test Network web site (read 7.1.2020), https://5gtn.fi/

[11] TR 21.915 V1.1.0 (2019-03), 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Release 15 Description; Summary of Rel-15 Work Items (Release 15), 3GPP, Sophia Antipolis, France, 112 p.

[12] 3GPP TR 22.891 V14.2.0 (2016-09), 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Feasibility Study on New Services and Markets Technology Enablers; Stage 1 (Release 14), 3GPP, Sophia Antipolis, France, 95 p.

[13] Top 5 Use Cases web site (read 5.8.2019), SDX Central, https://www.sdxcentral.com/5g/definitions/top-5g-use-cases/

[14] 5G use cases and requirements (2016), Nokia White Paper, Nokia, Espoo, Finland, https://www.ramonmillan.com/documentos/bibliografia/5GUseCases_Nokia.pdf , 12p.

[15] 5G Architecture White Paper, View on 5G Architecture (2019), version 3.0, June 2019, 5G PPP Architecture Working Group, https://5g-ppp.eu/5g-ppp-architecture-public-consultation/, 166 p.

[16] What is 5G Network Slicing (read 8.8.2019), SDX Central, https://www.sdxcentral.com/5g/definitions/5g-network-slicing/

[17] Ordonez-Lucena J., Ameigeiras P., Lopez D., Ramos-Munoz J.J., Lorca J. & Folgueira J. (2017) Network Slicing for 5G with SDN/NFV: Concepts, Architectures, and Challenges, IEEE Communications Magazine, vol. 55, no. 5, May 2017, p. 80-87

[18] 5G Americas White Paper – Network Slicing for 5G and Beyond (2016), 5G Americas, http://www.5gamericas.org/files/3214/7975/0104/5G_Americas_Network_Slicing_11.21_Final.pdf, 35 p.

[19] What is Software Defined Networking (SDN)? Definition (read 14.8.2019), SDX Central, https://www.sdxcentral.com/networking/sdn/definitions/what-the-definition-of-software-defined-networking-sdn/

[20] Jain R. & Paul S. (2013), Network Virtualization and Software Defined Networking for Cloud Computing: A Survey, IEEE Communications Magazine, vol. 51, no. 11, November 2013, p. 24-31

[21] What are SDN Northbound APIs (and SDN Rest APIs)? (read 16.8.2019), SDX Central, https://www.sdxcentral.com/networking/sdn/definitions/north-bound-interfaces-api/

[22] Han B., Gopalakrishnan V., Ji L. & Lee S. (2015), Network Function Virtualization: Challenges and Opportunities for Innovations, IEEE Communications Magazine, vol. 53, no. 2, February 2015, p. 90-97

[23] What is NFV – Network Functions Virtualization – Definition? (read 22.8.2019), SDX Central, https://www.sdxcentral.com/networking/nfv/definitions/whats-network-functions-virtualization-nfv/

[24] Ye Z., Cao X., Wang J., Yu H. & Yu H. (2016), Joint topology design and mapping of service function chains for efficient, scalable, and reliable network functions virtualization, IEEE Network, vol. 30, issue 3, May-June 2016, p. 81-87

[25] Bari F., Chowdhury S R, Ahmed R. & Boutaba R. (2015), On orchestrating virtual network functions. In: 2015 11th International Conference on Network and Service Management (CNSM), November 9-13 2015, Barcelona, Spain, DOI: 10.1109/CNSM.2015.7367338

[26] Taleb T.; Bagaa M. & Ksentini A. (2015), User mobility-aware Virtual Network Function placement for Virtual 5G Network Infrastructure. In: 2015 IEEE International Conference on Communications (ICC), June 8-12 2015, London, UK, DOI: 10.1109/ICC.2015.7248929

[27] Firmin F., The Evolved Packet Core web pages (read 30.9.2019), 3GPP, https://www.3gpp.org/technologies/keywords-acronyms/100-the-evolved-packet-core

[28] Ahmadi S., LTE-Advanced, A Practical Systems Approach to Understanding 3GPP LTE Releases 10 and 11 Radio Access Technologies (2014), Academic Press, DOI: https://doi.org/10.1016/C2012-0-02224-7 ,1152 p.

[29] Road to 5G: Introduction and Migration (2018) (read 26.8.2019), GSMA, https://www.gsma.com/futurenetworks/wp-content/uploads/2018/04/Road-to-5G-Introduction-and-Migration_FINAL.pdf.

[30] Parkvall S., Dahlman E., Furuskar A. & Frenne M. (2017), NR: The New 5G Radio Access Technology, IEEE Communications Standards Magazine, vol. 1, issue 4, December 2017, p. 24-30

[31] I C-L., Han S., Xu Z, Wang S., Sun Q. & Chen Y. (2016), New Paradigm of 5G Wireless Internet, IEEE Journal on Selected Areas in Communication, vol. 34, issue 3, March 2016, p. 474-482

[32] Kekki S. & others (2018), ETSI White Paper No. 28, MEC in 5G networks, First edition – June 2018, ETSI, 06921 Sophia Antipolis CEDEX, France, ISBN No. 979-10-92620-22-1, 28 p.

[33] Giust Fabio & others (2018), ETSI White paper No. 24, MEC Deployment in 4G and Evolution Towards 5G, First edition – February 2018, ETSI, 06921 Sophia Antipolis CEDEX, France, ISBN No. 979-10-92620-18-4, 24 p.

[34] Multi-access Edge Computing (MEC); Framework and Reference Architecture (2019), ETSI GS MEC 003 v2.1.1 (2019-01), ETSI, 650 Route des Lucioles, F-06921 Sophia Antipolis Cedex, France, 21 p.

[35] Dryjanski M (2018), IP and Mobile Trends and Education, 5G core Network Functions web page (Read 17.1.2020), https://www.grandmetric.com/2018/03/02/5g-core-network-functions/

[36] IMT vision – Framework and overall objectives of the future development of IMT for 2020 and beyond (2015), Recommendation ITU-R M.2083-0 (09/2015), Electronic publication, https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf , International Telecommunication Union, Radiocommunication Sector of ITU (ITU-R), Geneva, 2015, 21 p.

[37] Carugi M. (2018), Key features and requirements of 5G/IMT-2020 networks, https://www.itu.int/en/ITU-D/Regional-presence/ArabStates/Documents/events/2018/RDF/Workshop%20Presentations/Session1/5G-%20IMT2020-presentation-Marco-Carugi-final-reduced.pdf, ITU Arab Forum on Emerging Technologies, Algiers – Algeria, 14-15 Feb. 2018, 26 p.

[38] 3GPP TR 38.913 V15.0.0 (2018-06), 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on Scenarios and Requirements for Next Generation Access Technologies; (Release 15), 3GPP, 650 Routes des Lucioles, Sophia Antipolis, Valbonne, France, 39 p.

[39] 5GKPIs/5G Key Performance Indicators web page (read 1.10.2019), https://www.rfwireless-world.com/Terminology/5G-KPIs-Key-Performance-Indicators.html, RF Wireless World

[40] Euro-5g – Supporting the European 5G Initiative, D2.6 Final report on programme progress and KPIs, Version: v0.2 (2014), https://5g-ppp.eu/wp-content/uploads/2017/10/Euro-5G-D2.6_Final-report-on-programme-progress-and-KPIs.pdf , EURO-5G Consortium Parties, 33 p.

[41] Vuzix M300XL Smart Glasses Technical Specifications (2019), https://vuzix-website.s3.amazonaws.com/files/Content/docs/north-american/web/Vuzix-M300XL-Smart-Glasses-d01.pdf, web page read 5.11.2019), Vuzix, West Henrietta, New York, USA, 2 p.

[42] Kaitotek Qosium web pages (read 6.11.2019), Kaitotek, https://www.kaitotek.com/qosium

[43] The Linux PTP Project web pages (read 19.11.2019), http://linuxptp.sourceforge.net/

10 APPENDICES

Appendix 1 Test case specification

Appendix 1 Test case specification

Test case specification

Test designed by: Antti Pauanne
Test designed date: 16.12.2019
Test executed by:
Test execution date:

Pre condition:
Hospital Use Case Streaming system is functional but not streaming
Qosium measurement system installation done including probes installation

Dependencies:
Oulu University 5G Test Network functional

Scope
Purpose is to measure data transmission delays

1. From streaming PC connected to 5G TN to NGINX multimedia server residing in Edge Server in the 5G Test Network. (Streaming to NGINX.)
2. From NGINX multimedia Server in the 5G TN Edge Server to PC receiving the stream connected to 5G TN. (Streaming from NGINX

Test steps for NGINX at Edge Server. 5G.

1. Send stream from the Operating Room to NGINX.
    a. Use 5G network
    b. Limit bandwidth of the stream to 500kb/s in OBS
    c. Do not limit the duration of the stream but keep it up indefinitely
    d. Verify from OBS that NGINX is receiving the stream
2. Receive stream from the NGINX in the Observation room
    a. Use 5G network
    b. Verify from VLC that stream is visible in the screen
3. Start Qosium Scope measurement
    a. Activate sending PC probe
    b. Activate NGINX probe
    c. Activate receiving PC probe
    d. Start Qosium measurement for sending the stream
        i. Select delay to be recorded and displayed
        ii. Keep the measurements going on for 5 minutes
        iii. Stop the measurement
        iv. Store the measurement data using filename Edge_5G_send_delay_5min.
        v. Start the measurement again and now keep them going for 1 minute
        vi. Store the measurement data using filename Edge_5G_send_delay_1min.
    e. Start Qosium measurement for receiving the stream
        i. Select delay to be recorded and displayed

ii. Keep the measurements going on for 5 minutes
iii. Stop the measurement
iv. Store the measurement data using filename Edge_5G_receive_delay_5min.
v. Start the measurement again and now keep them going for 1 minute
vi. Stop the measurement
vii. Store the measurement data using filename Edge_5G_receive_delay_1min.

Test steps for NGINX at Edge Server. 4G.

1. Send stream from the Operating Room to NGINX.
    a. Use 4G network
    b. Limit bandwidth of the stream to 500kb/s in OBS
    c. Do not limit the duration of the stream but keep it up indefinitely
    d. Verify from OBS that NGINX is receiving the stream
2. Receive stream from the NGINX in the Observation room
    a. Use 4G network
    b. Verify from VLC that stream is visible in the screen
3. Start Qosium Scope measurement
    a. Activate sending PC probe
    b. Activate NGINX probe
    c. Activate receiving PC probe
    d. Start Qosium measurement for sending the stream
        i. Select delay to be recorded and displayed
        ii. Keep the measurements going on for 5 minutes
        iii. Stop the measurement
        iv. Store the measurement data using filename Edge_4G_send_delay_5min.
        v. Start the measurement again and now keep them going for 1 minute
        vi. Stop the measurement
        vii. Store the measurement data using filename Edge_4G_send_delay_1min
    e. Start Qosium measurement for receiving the stream
        i. Select delay to be recorded and displayed
        ii. Keep the measurements going on for 30 minutes
        iii. Stop the measurement
        iv. Store the measurement data using filename Edge_4G_receive_delay_5min.
        v. Start the measurement again and now keep them going for 1 minute.
        vi. Stop the measurement
        vii. Store the measurement data using filename Edge_4G_receive_delay_1min.

Test steps for NGINX at Cloud Server. 5G.

1. Send stream from the Operating Room to NGINX.

   a. Use 5G network
   b. Limit bandwidth of the stream to 500kb/s in OBS
   c. Do not limit the duration of the stream but keep it up indefinitely
   d. Verify from OBS that NGINX is receiving the stream
2. Receive stream from the NGINX in the Observation room
   a. Use 5G network
   b. Verify from VLC that stream is visible in the screen
3. Start Qosium Scope measurement
   a. Activate sending PC probe
   b. Activate NGINX probe
   c. Activate receiving PC probe
   d. Start Qosium measurement for sending the stream
      i. Select delay to be recorded and displayed
      ii. Keep the measurements going on for 5 minutes
      iii. Stop the measurement
      iv. Store the measurement data using filename Cloud_5G_send_delay_5min.
      v. Start the measurement again and now keep them going for 1 minute
      vi. Stop the measurement
      vii. Store the measurement data using filename Cloud_5G_send_delay_5min.
   e. Start Qosium measurement for receiving the stream
      i. Select delay to be recorded and displayed
      ii. Keep the measurements going on for 5 minutes
      iii. Stop the measurement
      iv. Store the measurement data using filename Cloud_5G_receive_delay_5min.
      v. Start the measurement again and now keep them going for 1 minute
      vi. Stop the measurement
      vii. Store the measurement data using filename Cloud_5G_receive_delay_1min.

Test steps for NGINX at Cloud Server. 4G.

1. Send stream from the Operating Room to NGINX.
   a. Use 4G network
   b. Limit bandwidth of the stream to 500kb/s in OBS
   c. Do not limit the duration of the stream but keep it up indefinitely
   d. Verify from OBS that NGINX is receiving the stream
2. Receive stream from the NGINX in the Observation room
   a. Use 4G network
   b. Verify from VLC that stream is visible in the screen
3. Start Qosium Scope measurement
   a. Activate sending PC probe
   b. Activate NGINX probe
   c. Activate receiving PC probe
   d. Start Qosium measurement for sending the stream

i. Select delay to be recorded and displayed
ii. Keep the measurements going on for 5 minutes
iii. Stop the measurement
iv. Store the measurement data using filename Cloud_4G_send_delay_5min.
v. Start the measurement again and now keep them going for 1 minute
vi. Stop the measurement
vii. Store the measurement data using filename Cloud_4G_send_delay_1min.

e. Start Qosium measurement for receiving the stream
   i. Select delay to be recorded and displayed
   ii. Keep the measurements going on for 5 minutes
   iii. Stop the measurement
   iv. Store the measurement data using filename Cloud_4G_receive_delay_5min.
   v. Start the measurement again and now keep them going for 1 minute
   vi. Stop the measurement
   vii. Store the measurement data using filename Cloud_4G_receive_delay_1min.