IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

5G Network Architecture - Design Considerations

SUHAS U
ROHAN H RAJ
TANISH C S

BANGALORE INSTITUTE OF TECHNOLOGY, BANGALORE-560004.

Abstract

The data rates of up to 10 GB/s will characterize 5G networks telecommunications standards that are envisioned to replace the current 4G/IMT standards. The number of network-connected devices is expected to be 7 trillion by the end of this year and the traffic is expected to rise by an order of magnitude in the next 8 years. It is expected that elements of 5G will be rolled out by early 2020s to meet business and consumer demands as well as requirements of the Internet of Things. [1] China's Ministry of Industry and Information Technology announced in September 2016 that the government-led 5G Phase-1 tests of key wireless technologies for future 5G networks were completed with satisfactory results. This paper presents an overview of the challenges facing 5G before it can be implemented to meet expected requirements of capacity, data rates, reduced latency, connectivity to a massive number of devices, reduced cost and energy, and appropriate quality of experience.

Introduction

5 G technology provides a new mobile network architecture that, beyond a simple increase in network speed, represents a new paradigm in the end-to-end design of the network and the deployment of new services.[4] The 5G system (5GS) is addressed globally, covering new frequency bands, new radio access technologies, new core capabilities as well as the overall optimization from the mobile radio to the backhaul and core network to run innovative services [4]. Enabled by 5G, the telecom industry is moving beyond voice and mobile broadband, the two core services that have powered the industry until today. Two main opportunities for telecom services involve the "massive" connection of machinetype objects and the "ultra-reliable and low latency" services, in addition to an enhanced mobile broadband service with even higher throughput than 4G. This new service offering entails a transition from a "horizontal" service delivery model, where services are defined independently from the consumers' needs, towards a "vertical " delivery model, where the provided services are tailored to specific industry sectors and verticals [4]. The transition to the "vertical" delivery model is expected to have a strong impact on the underlying business model of mobile networks. The use cases enabled by 5G address several industry segments that were not effectively covered by 4G and can potentially provide mobile network operators

Mobile communications have changed from a system capable of transmitting voice conversations for millions of users to a system supporting trillions of devices, primarily transporting data rather than voice [1]. The needs for data communications are different than that of voice, and the current system of 4G/Long-Term Evolution (LTE) networks will be at maximum capacity in the next few years. New technologies and new network architectures will be required to meet the needs for future users – and there will be many of them (e.g. Andrews et al., 2014). Wireless World Research Forum predicts that there will be over 7 trillion network-connected wireless devices by the end of 2017. (Wang 2014, 123) Experts predict that traffic will be 1,000 times higher than today by 2025. (Dahlman 2014, 43) This will drive the need for higher data rates yet reduced energy consumption. [3] These goals cannot be accomplished using the current 4G technology and architecture. Six main challenges are required to be solved in order for 5G to meet expected requirements, and many new technologies and architecture-types are believed to be able to accomplish that.

Limitations of 4G/Challenges for 5G

The vast majority of the speed and capacity gains made in the last decade were through improvements in the efficiency of air-spectrum use (such as modulation and coding schemes) and through spectrum acquisition. (Li 2014) These methods are nearing their limit, so other methods for increase will have to be explored. Current data rates are 1 GB/s for slow or stationary devices, or 100 Mb/s for mobile devices.[1] 5G seeks out a 10 GB/s data rate. (Wang 2014, 123) Again, technology will have to change in order to meet that demand. Base stations account for 70 percent of cellular operators' electricity usage. Increasing capability would only increase that percentage, which is unsustainable.

4G is considered to be a descendant of 2G and 3G. Like the evolutions between earlier generations, 4G took elements of 3G and made them better.[1] Instead of maintaining both circuit switching and Internet Protocol (IP) packets, 4G moved to IP for all services. Advanced radio technologies such as orthogonal frequency-division multiplexing (OFDM) and multiple-input multiple-output (MIMO) antenna arrays were incorporated to make better use of the current spectrum. (Wang 2014, 122) Better Quadrature Amplitude Modulation (QAM) techniques also increased data throughput. These advances brought data transmission speeds into the range high enough for multimedia, especially streaming video. Now that video service is available, mobile broadband use skyrocketed. That led to increasing smartphone usage, and many devices began to be developed. Large format, high-resolution screens nearly always-on and always-communicating with base stations has led to high power usage with battery technologies that have met their limits.

It is clear that the many small changes made to advance through the previous generations will not be enough to make the jump to 5G requirements.[1] Passage from first generation through 2G, 3G, and 4G may have been evolutions on the same theme, but many of the techniques that have increased capabilities up to this point are nearing their limits. It seems that a revolution in the mobile communications system must take place. Drastic changes to the entire architecture of cellular networks will have to be made.

Main Challenges

The standards for 5G have not yet been created, with future approval expected to occur in 2020. However, telecommunications companies and the International Telecommunication Union have been working on requirements for the next generation since the last standardization of the International Mobile Telecommunications – Advanced standard in 2008, but specifically for 5G since 2015.[1] (3GPP 2015) Much of the problem creating the need for a new generation of mobile telecommunications comes from a wide expansion of use cases. What once only needed to transmit voice grew to data. And what was once just a handheld device is growing into many disparate devices, many of which have no human interaction whatsoever. These new use cases, which will be detailed later in this paper, have created a great number of challenges that the current system cannot support, and simple expansion of the current technologies do not seem to be a solution for. The goals of 5G are to support these many new use cases.

More Capacity

5G network architects predict the need to support a 1000-fold increase in traffic compared to 2010.
[1] Mobile traffic grew nearly 70% during 2013, reaching 2.5 exabytes a month. They expect it to grow to 10 times that monthly amount by 2019. Theodore Rappaport of the New York University Wireless research center says 4G "can never accommodate this new demand." (Young 2015) 4G has gotten to where it is capacity- 3 wise through radio technology advances and acquisition of new spectrum (Figure 1.) These methods are meeting their limits. Most modulation techniques and coding schemes cannot get more efficient, and the spectra best suited for long distance communications has been allocated.

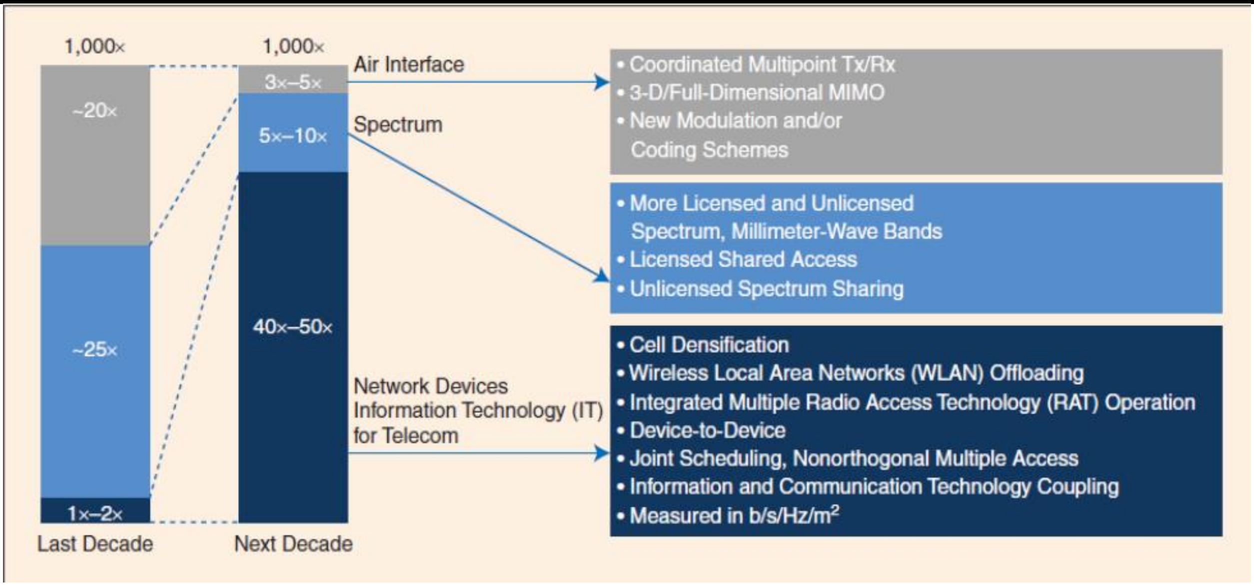


Figure 1 - Network Capacity Growth (Li 2014, 72)

Higher Data Rate

5G goals include data rates at least ten times higher than current peak data rates. Current data rates are limited for the same reasons that capacity is limited.[1] Increased usage of the same frequency space limits the amount of bandwidth that can be used per device. Additionally, the current capabilities of backbone, backhaul, and fronthaul limit the amount of data that can pass from base stations to mobile devices and vice versa. (Agyapong 2014, 67) Solutions will require expansion beyond current means rather than more efficient usage of the current means to affect data rate.

Lower Latency

The Internet of Things (IoT) is putting a higher demand on always-on and always–available abilities of wireless networks. Connected devices in traffic control, autonomous cars, healthcare, and industrial automation cannot be subject to latency Figure 1 - Network Capacity Growth (Li 2014, 72) 4 issues when lives may be on the line. This extremely-high-reliability and extremely - low latency requirement is not supportable by the current model.[1] (Agyapong 2014, 66) Previous generations of mobile telecommunications and their users primarily used static or latency-independent data, such as websites and email. Higher data rates led then to streaming video, which can be affected by latency, though buffering can prevent that. But the expansion of real-time applications, such as teleconferencing and time-critical communications, such as remote controlled medical robots, cannot suffer the effects of latency to properly function.

Connectivity for Massive Number of Devices

Not only are the numbers of devices exploding due to IoT, but the different connection needs makes this a wide-ranging challenge. [1] No longer are the needs expected to be similar to previous requirements (voice calls, <150 ms latency acceptable.) Devices need to be able to range from sleeping until needed to always-on, yet get the connectivity when required. (Agyapong 2014, 68) Handling the massive number of devices that may pop in and out of connection constantly will be a challenge. 4G is not ideal to handle this because current systems are disconnected, according to Zhiguo Ding of Lancaster University. (Hellemans 2015) Bluetooth, RFID, and other various short-distance communication protocols are not set up to communicate with each other. There is no common system. The challenge will be creating a system that is common yet able to support a myriad number of devices with hundreds of different use cases.

Data Analytics and AI/ML-Based Network Optimization:

Data collection and analysis are of primary importance in 5G networks (both public and private) due to the increasing need for efficient monitoring and management of the network and underlying infrastructures. At the same time, technologies like data analytics and AI/ML are investigated as a means to facilitate network automation and result in enhanced utilization of the network resources [5]. Network automation is particularly important for the cost-efficient deployment of NPNs since it has the potential to significantly reduce management costs. Towards this direction, AI/ML techniques are of the utmost importance for the deployment of an intelligent, cognitive, and flexible 5G system, primarily due to the memorization and accurate decision-making capabilities [5]. As a matter of fact, ML algorithms have already been proven to represent a powerful tool with many potential applications to enhance and facilitate wireless communication networks in several areas, such as radio channel modeling, channel estimation, signal detection, network management and performance improvement, access control, and resource allocation. In our proposed architectural approach, as will be described in the following section, the system incorporates modules that are in charge of gathering information from all architectural layers and provide this information either to the registered services or to AI/ML-based modules for network and service optimization. In this context, cross-layer data gathering is supported. In addition, network and management related data are processed by AI/ML algorithms in order to optimize specific parameters and key performance indicators (KPIs). Typical examples include network configuration (RAN planning and deployment, scheduling), cross-slice optimization for improved quality of service (QoS), energy consumption by each UE, as well as mobility management.

Reduced Cost and Energy

Energy storage capabilities are not keeping up with the media abilities of current devices, and the desire to increase those abilities will require a reduction in the amount of energy required in order to communicate with the network.[1] The IoT also is affecting this issue because small, low-power devices such as sensors are expected to run on a battery for several years. (Dahlman 2014, 42) On the provider's side, base station electricity usage and the cost of equipment upgrades to handle the new requirements is a concern that needs to be solved for the business to remain profitable. Current estimates of the percentage of radio network power usage are 70 to 80 percent of total operations energy usage. (Agyapong 2014, 68) This is driving costs higher, and future usage will only increase that. Systems will need to change in order to meet service demands yet maintain reasonable costs.

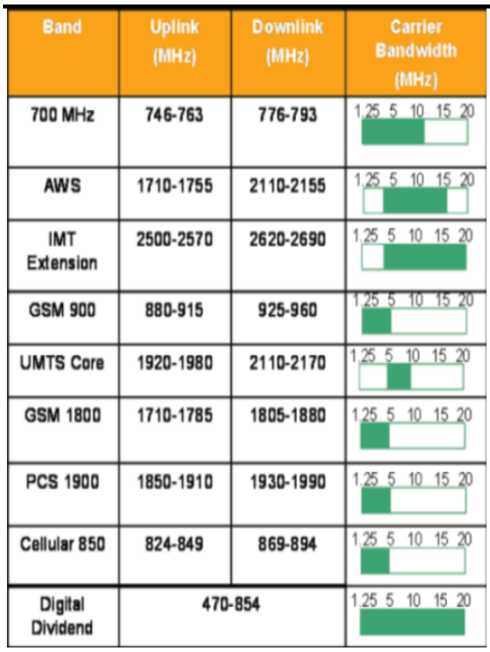
Quality of Experience

Quality of Experience (QoE) is the user's perception of how the experience of using a device meets expectations. It is related to some of the above challenges, such as data rate and latency, but is combined in such a way that it creates a balance between them. High data rate and low latency may make a streaming video look good, but it is a large drain on the battery.[1] However, a low QoE will cause a user to become dissatisfied. Due to 4G architecture, it is difficult to maintain consistent experience at all times and in all locations.

5G Design Proposal Elements

Many proposals have been produced to improve the technology and architecture of the cellular network in order to achieve the goals of 5G capabilities.[1] Though the proposals vary, many suggested elements are common across multiple papers and are possible solutions to the primary goals of 5G. These solutions, combined to work in coordination with each other, have the potential to meet the demands of mobile wireless usage for the next decade. Individual solutions, as well, are even able to meet several goals and overcome multiple challenges on their own. However, many of the proposed changes to the telecommunications system will drive additional modifications to how the network is constructed or functions in order for them to work. They need to work in concert with the other proposals in order to fully meet the expected requirements for 5G telecommunications.5G Design Proposal Elements

Many proposals have been produced to improve the technology and architecture of the cellular network in order to achieve the goals of 5G capabilities. Though the proposals vary, many suggested elements are common across multiple papers and are possible solutions to the primary goals of 5G. These solutions, combined to work in coordination with each other, have the potential to meet the demands of mobile wireless usage for the next decade. Individual solutions, as well, are even able to meet several goals and overcome multiple challenges on their own. However, many of the proposed changes to the telecommunications system will drive additional modifications to how the network is constructed or functions in order for them to work. They need to work in concert with the other proposals in order to fully meet the expected requirements for 5G telecommunications.



Band	Uplink (MHz)	Downlink (MHz)	Carrier Bandwidth (MHz)
700 MHz	746-763	776-793	1.25 5 10 15 20
AWS	1710-1755	2110-2155	1.25 5 10 15 20
IMT Extension	2500-2570	2620-2690	1.25 5 10 15 20
GSM 900	880-915	925-960	1.25 5 10 15 20
UMTS Core	1920-1980	2110-2170	1.25 5 10 15 20
GSM 1800	1710-1785	1805-1880	1.25 5 10 15 20
PCS 1900	1850-1910	1930-1990	1.25 5 10 15 20
Cellular 850	824-849	869-894	1.25 5 10 15 20
Digital Dividend	470-854		1.25 5 10 15 20

Figure 3 - Current 2G, 3G, 4G, & LTE-A spectrum and bandwidth allocation (T. Rappaport 2002)

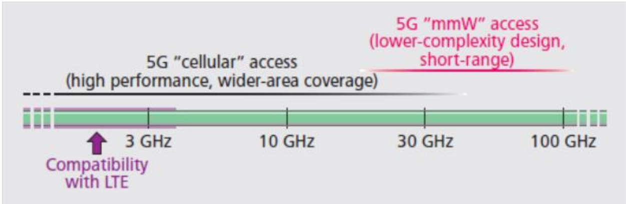


Figure 2 - Spectrum range to be considered for 5G wireless access (Dahlman 2014, 45)

Spectra

Much of the spectrum below a few gigahertz is already heavily populated. These frequencies are allocated, and mostly licensed, for other uses (radiolocation, satellites, radio and television broadcast, etc.) The current range for mobile telecommunications is between 700 MHz and 2.7 GHz. (Figure 2) Expansion into higher frequencies (>3GHz), and even millimeter wave frequencies (>30GHz) will be required simply because there is no more room for lower frequency allocation. (Figure 3)

However, these higher frequencies have problems with signal propagation over large distances. There are several factors involved in this attenuation. First, signals naturally attenuate over distances traveled through over-the-air transmission. This attenuation is proportional to the square of the frequency of the signal. Therefore, entering into these higher frequency ranges will lower the range, assuming transmission power remains constant. But additional factors make these frequencies even worse. Common atmospheric gasses, such as oxygen and water vapor, strongly absorb signals of specific frequencies. Rainfall is an even greater obstruction than the gasses, causing nearly 10 db/km attenuation at all super high frequencies and above. Most damaging to frequency propagation are solid materials. Even traveling short distances, gigahertz signals lose significant power traveling through materials such as brick, tinted windows, interior walls and cubicles. Power losses could easily be over 40 dB of attenuation. (T. S. Rappaport 2013, 340)

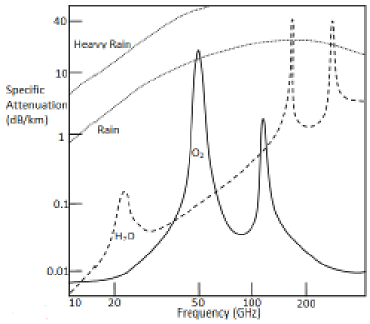


Figure 4 - Atmospheric Attenuation - modified from (Radiocommunications Div., U.K. Dept. Trade Industry 1988) and (T. S. Rappaport 2013)

Environment	Material	Thickness (cm)	Penetration Loss (dB)
Outdoor	Tinted Glass	3.8	40.1
	Brick	185.4	28.3
Indoor	Clear Glass	<1.3	3.9
	Tinted Glass	<1.3	24.5
	Wall	38.1	6.8
	3 Walls	N/A	41.1
	3 Walls, 2 Doors	N/A	Weak Signal Detected
	5 Walls, 2 Cubicles, 1 Elevator	N/A	No signal Detected

Table 1 – Penetration losses at 28 GHz – modified from (T. S. Rappaport 2013)

Massive Multiple-Input Multiple-Output

Multiple-Input multiple-output (MIMO) antenna arrays are currently in use, but[1] generally only have a few antennas. MIMO arrays contain multiple antennas which can be used to beam shape and direct transmitted signals (and obtain received signals) in a specific direction. This technology can increase the channel efficiency, increasing data rates as well as lowering energy usage as lower-power signals can be transmitted and received. At current mobile frequencies, wavelengths are on the order of about 1 foot long. So the half-wavelength size of the antenna is about 6 inches. There simply isn't much room for more than a few antennas. However, if the super high frequencies and millimeter wave frequencies proposed for 5G begin to be utilized, antennas could be on the order of a centimeter or less. Many more antennas could be packed easily inside of handheld mobile devices than could ever fit before.

[1]Massive MIMO (mMIMO) will have tens or hundreds of antennas and not only at the base station but dispersed throughout the cell. These multi-antenna arrays will be placed in well-planned, line-of-sight locations to best utilize their properties. They can be placed on buildings, connected to their internal wireless access points, which will be discussed in a later section. (Gupta, 1208) The antenna arrays can steer beams toward a connected device, reducing power needed to communicate while also producing a low interference signal. The beams can also be steered into using the reflective properties of millimeter waves to circumvent the limitations of the normally line-of-sight signals. (T. e. Rappaport 2014, 111) Signals can intentionally be directed into buildings and other solid objects in order to reflect their signal into an area that the antenna cannot reach directly due to physical obstructions. These directed signals provide a strong link that is not subject to as much signal fading as current mobile communications systems experience. Currently, when a device is in a location of obstruction, multipath signals can then enter from the base station and act destructively, and the device must wait for a different transmission channel to send on. This is known as fading, and it is the primary cause of latency in telecommunications.

[3] Massive MIMO can be used to directly meet the goals of 5G by increasing the data rate and decreasing latency. It also indirectly meets the goals by allowing the millimeter wave frequencies to be practically useful despite their characteristics of high attenuation and low diffraction.

Device-to-Device Communication

Device-to-device communication (D2D) is the method of devices transmitting and receiving data between each other on the user plane without having to utilize network resources. [1] Conventional communications has all devices communicate directly with – and only with – the base station.) D2D would allow multiple scenarios in which devices communicate with each other and the base station, sometimes extending cell range with an ad hoc network of hops from device to device. Devices communicating directly with each other will reduce the energy required to communicate due to shorter distances traveled and less communication due to the direct path utilized instead of relaying through the base station first.

D2D can also free frequencies and base station overhead if it is not communicating with the base station for all communication, thus increasing capacity in the cell. D2D can improve data rate and lowers latency.

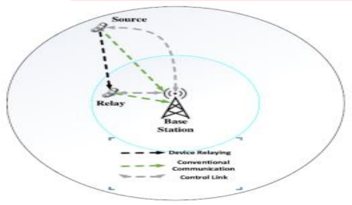


Figure 5 - Device relaying communication with base station controlled link formation. (Gupta 2015, 1218)

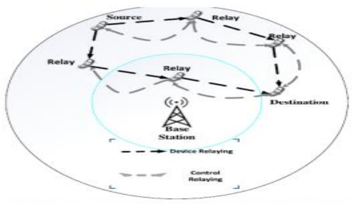


Figure 6 - Device relaying communication with device controlled link formation. (Gupta 2015, 1219)

Data can even be sent across multiple devices, with the middle devices acting as relays between the source and destination devices. D2D can also reutilize spectrum, increasing efficiency. This relaying-between-devices has the added benefit of again making up for the diffraction deficiencies of millimeter waves. Devices obstructed from line-of-sight view of an antenna can have their signals relayed through a device that is in LOS with an antenna, ensuring a good connection no matter the location.

D2D communications can contribute greatly toward meeting the goals of 5G. Lower base station usage saves energy, frees frequencies for more capacity, decreases latency, increases data rate, and like mMIMO, helps to lessen the limitations of millimeter wave LOS issues. All of these gains work together to meet user quality of experience.

Reduced Protocol Overhead

Reducing overhead in nearly any scenario is a means to gain efficiency. Even with the simplest of mobile data transmissions, many functions take place behind the scenes that ensure that the data is sent correctly.[1] Outside of even data transmissions, control packets are sent back and forth between base stations and devices regularly. These packets often have no payload, yet account for 46% of transmitted packets. Compressing these packets could double the efficiency of the network. (Li 2014, 75) Reducing even the amount of these packets being sent would be beneficial especially to IoT devices. (Dahlman 2014, 43) Small sensors, often running on batteries, need to conserve as much energy as possible. Constant chatter with the base station simply transmitting control packets would quickly deplete energy sources. Being able to configure how and when devices communicate with the base station is vital for small IoT devices.

D2D communications will reduce communications between devices and base stations and data centers and all of their associated overhead will be reduced with it. Multipath Transmission Control Protocol (TCP) can be incorporated to allow devices to utilize whatever connection is available, without having to create new paths constantly. Simplifying protocols can also reduce the amount of data having to travel on backhaul, thereby increasing data rates.

Reducing the protocol overhead meets 5G goals by increasing data rate, reducing power consumption, and supports the IoT especially for low-power devices.

Network Function Virtualization

A software-defined network solution is also proposed to increase efficiency as well as reduce costs. [1] Currently, many of the functions of control and network management are located at the base stations. Upgrades or changes to functions are difficult and costly to implement. By removing specialized hardware at the base stations and virtualizing them in the cloud (Figure 9), upgrades can be made more frequently and at lower cost. Management of the entire network can be handled more efficiently at a data center, allowing downline base stations to handle only what is required at their end. Analysis at the data center level can also include network intelligence to better route data and manage connections.

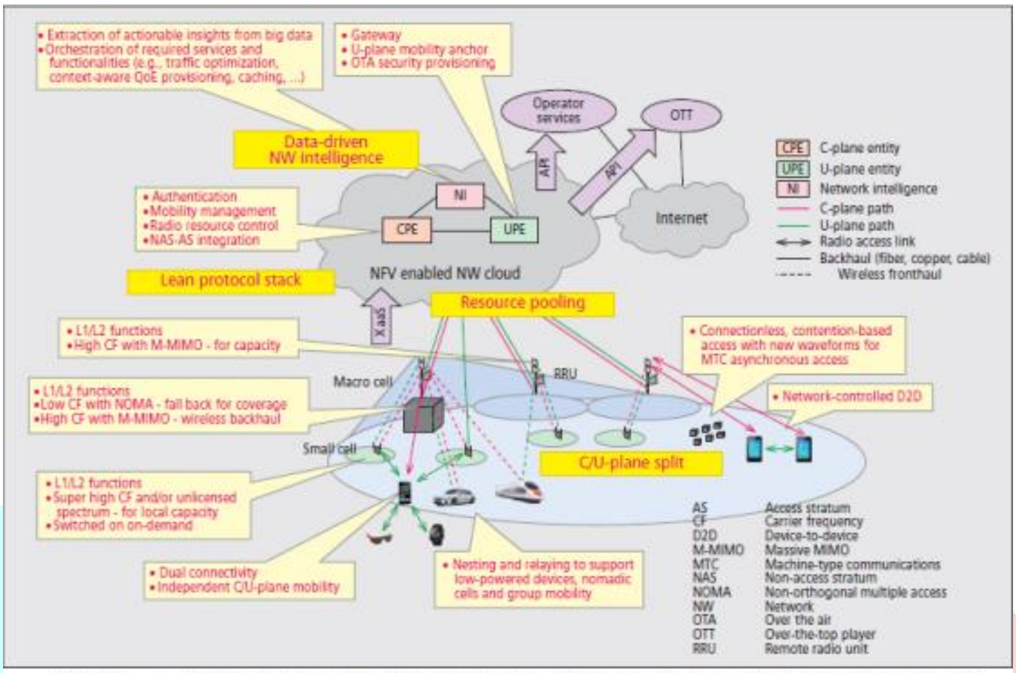


Figure 9 – Network Architecture highlighting cloud-based management (Agyapong 2014, 72)

Passing devices from base station to base station as they move between cells would be much simpler. Currently, base stations take care of that between each other, which updates their own databases and other central office databases. If controlled at the core network, it could free base stations of that responsibility and lessen the overhead involved. Deep Packet Inspection is a tool used to see the type of data being transmitted. By knowing the contents of packets, the network can better route the data by altering connections so that the user's quality of experience can be increased. This sort of function needs to be done at a data center. It is too late in the flow of data for it to be done at the base station, so the quality of a streaming video, for instance, cannot be increased at the base station level.
Network Function Virtualization supports the 5G goals by decreasing costs and increasing user quality of experience.

KEY WORKING PRINCIPLES :

Wireless Backhaul Integration The most important feature of the Frugal 5G network is the integration of MMN in the access network [2]. Typically, the cellular standards support relays which enable not more than two hops in the network. We envision the MMN to be a multihop mesh network. Therefore, integration of MMN within the Frugal 5G network is essential so as to deliver the required QoS to the end-users. When a data packet has to be delivered to a user through the WLAN, the flow controller gives the required instructions to both the WLAN and the middle mile controllers and the required path is set up through the MMN and the WLAN. Note that, the MMN provides infrastructure for WLAN APs to communicate with the WLAN controller as well and the control information exchange between the WLAN controller and WLAN APs takes place over the same MMN, which carries the data for users.

Flexible Fog Control :

The usage of NFV in AN (Fog) as well as CN allows for a flexible instantiation of functions across Fog and Cloud. [2] The SDN Controller in the Cloud provides the policy configuration to the management and orchestration entity for instantiation of different network functions on the individual Fog elements. The AN architecture as shown in Fig 4 is one of the possible ways in which network functions can be instantiated. Their instantiation in the Fog may vary and is dependent on the availability of the resources. For instance, when the compute and storage resources are limited, then only a handful of network functions, e.g., radio specific data plane functions, forwarding functions, Core Network Interworking Function etc. are instantiated on a Fog element. On the other hand, if compute and storage resources are available but the backhaul bandwidth (network resources) is limited, a larger number of functions can be installed in the Fog, especially the functions which help reduce the backhaul usage, e.g., the Content Server.

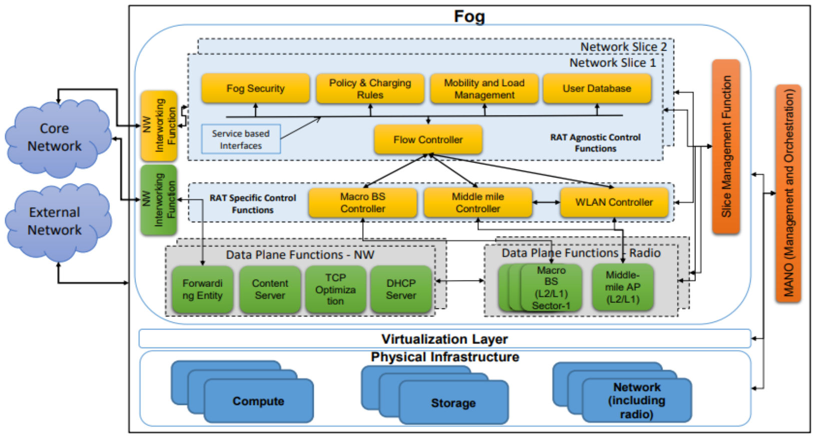


Fig. 4: Detailed architecture of Access Network (Fog)

Multi-operator Resource Sharing :

Sharing network resources is an important technique that can considerably reduce the network cost. With the help of NFV and SDN, it is possible to split the physical network resources into multiple logical networks also called network slices. Each of these network slices or logical networks can be used for different purposes, e.g., the individual slices can be allocated to an operator through which the operator can serve the rural areas. It is important to understand how the network resources are split among multiple slices[2]. Depending on the various parameters such as subscriber base, an operator might need a certain amount of network resources. As already mentioned, the RSCFs provide an abstract (virtual) view of the underlying network resources to the SMF. The SMF splits these virtual network resources into multiple network slices, each of which is allocated to individual operators. The RACFs are instantiated separately for each of these slices as they may be specific to the operator's network. This provides complete flexibility in terms of network resource sharing. When one operator's network (slice) experiences a low load scenario, the resources can easily be diverted to another operator's network (slice), which may be experiencing higher load.

Localized Communication Support :

In order to enable localized communication in the AN, without the involvement of Cloud, we have instantiated the relevant network functions in the Fog as explained in Section IV. [2] This results in improvement of backhaul efficiency (by avoiding the Cloud-Fog signaling/data transfer) and user experience. To understand how the localized data paths are handled by Fog, it is important to understand the data/traffic flow set-up in the network. Various possible data/traffic flows in the network are illustrated in Fig 5. A group of SDN controllers in the ANs along with the CN is typically responsible for setting up the data path through the network for a user-specific traffic flow. Flow Controller in the AN sets up the data path through the data plane entities under its control. A data path for a user is selected either through a Macro BS or through the wireless MMN and a WLAN AP. Depending on various parameters, such as, the mobility of the user, the strength of the radio signal, QoS requirements, and load on different RATs, the flow controller selects an appropriate data path for a user. Now, it is possible that an end-to-end data path for a traffic flow is fully contained within a single Fog element. Examples for such paths may be a Voice over IP (VoIP) session between two mobile users, who are in the close vicinity of each other (say, in the same village) or a data session between a user and the Content (cache) Server in the same Fog element. Usage of such data paths reduces the end-to-end latency of the flow and optimizes the resource utilization in the network by not using the wireless backhaul and the core Cloud elements. Since Fog elements can continue to provide services of localized nature in the absence of connectivity to the cloud, the network resilience is improved. Therefore, we may reduce the cost of network equipment by employing limited fault tolerance in the equipment.

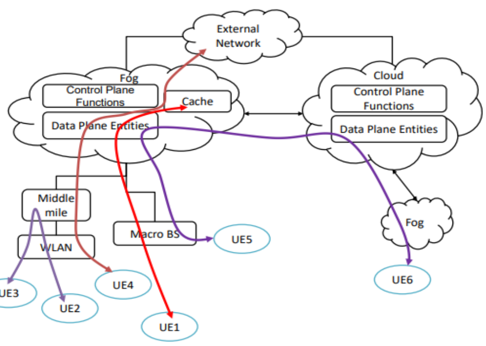


Fig. 5: Data Flow Examples

MAIN COMPONENTS :

This describes the primary building blocks of Frugal 5G Network Architecture as shown in Fig. 4.
Here categorize all network functions in AN as well as CN under control plane or data plane functions.[2] The CN is responsible for overall control of the Frugal 5G Network. The CN Control Plane performs the standard tasks, such as, mobility management, policy and charging control similar to the core network of other cellular systems. The CN sets up the traffic flow for a user through these data plane entities. They can also set up a path between two Fog elements, if required. Additionally, the CN control plane manages the distributed SDN controllers in each of the Fog elements. The data plane functions in the CN are responsible for forwarding the received data to another data plane entity in the network or to the external data network. The control and data plane functions across Fog and Cloud are synchronized with each other to avoid any inconsistencies. We now describe, in detail, the working of control and data planes in the AN.

Practical Considerations and Infrastructure Requirements

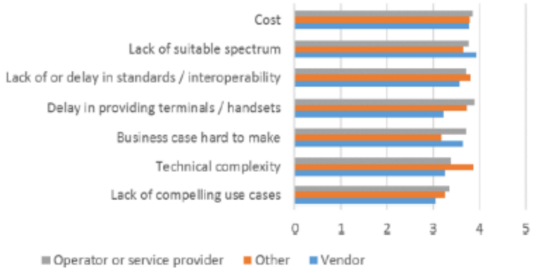


Figure 10 – Main barriers to 5G development and deployment (5 = very serious barrier; 1 = not a barrier at all) (n=97)

The many technological advances addressed in the previous section are viable solutions to the problems and challenges stated prior and may be able to meet the proposed requirements for 5G networks when considered all together.

However, turning these possible solutions into actual, implemented hardware and infrastructure will be a gargantuan task with many obstacles in the way. These types of concerns are rarely discussed in scholarly articles; however they are a critical concern in the deployment of mobile networks, being a consumer business, and factors greatly into the design of the system. In a Global mobile Suppliers Association survey of the industry on the challenges to meeting the 5G goals, most all surveyed scored the list of barriers at least at a 3, and most close to a 4. Cost, lack of suitable spectrum, and delay of standards/interoperability were the three highest ranked issues.

Cost

Previous generations of mobile technology certainly had increased costs when transitioning, but for the most part, the changes were minimal.[1] Hardware needed to be changed at the base stations and central office, but little more was done. The solutions considered in this paper will require an enormous increase in infrastructure costs. Network virtualization of functions will require large upgrades of the core servers to control the network. Changes will need to be made at every base station in order to accommodate the cloud-based computing and control of the towers. Perhaps most costly will be the additional small cell deployments. According to wirelessestimator.com, there are currently about 120,000 cell towers in the US.[1] Assuming only one additional small cell per current tower, that doubles the current number of towers. These new cells may not need full size towers, however, because they may be placed on buildings and will be located in more densely populated areas. But that does not negate the fact that the available real estate must be found, then purchased or leased, properly zoned, and power and backhaul utilities run to the locations.) shows how the current paradigm (the one large macro cell tower) is augmented with multiple antennas throughout the macrocell and femtocells on trains, buses, and buildings. There is also a reliance on femtocells in homes routing their wireless communications through hardwired internet connections. To get the speeds of 5G, houses will have to have fiber run to every home, which is usually only done today in brand new developments because ISPs do not find it cost effective to run it everywhere. Most homes have only copper connections. The goals of ultra-low latency may be great, but that is only on the radio network side. The overall latency of the user is a factor of the radio network, but also of all other connections on the internet to get the content desired. The cost to get low latency on the radio network, which is only part of the entire latency string, is not something most consumers would want to pay for.

Lack of Suitable Spectrum

Government entities that regulate the radio spectrum in their countries have control of what and how much of a frequency can be allocated to mobile telecommunications.[1] The process to open new spectra can sometimes be a slow, arduous task that does not keep up with the speed at which industry desires. It may also be done in a way that hampers the full and fair usage of that spectrum. In the US, the FCC recently opened several large sections of high frequency spectrum for 5G usage. However, some fear that it is not enough and that the big carriers will license portions that they will then not use completely, only covering large city centers and venues with 5G and leaving the rest of the country without. This locking down of frequencies will prevent true and full deployment of 5G in all areas.

Role of Satellites in the 5G networks :

The 5G system has seen increased demands on the back-haul with an increasing number of HetNets and small cells.[6] Satellites will play a major role in the extension of the 5G cellular network to new areas such as ships on the sea and remote land areas which are not covered by cellular networks. Also, high throughput Satellite Communication (SatCom) systems will be able to complement terrestrial provision in an area where it is difficult to do so with other terrestrial cells such as LTE. Indeed, integrating satellites in the future 5G networks will be seen as an essential part of the terrestrial infrastructure to provide strategic solutions for critical and lifesaving services. Satellites would be able to collect and distribute data from clusters of sensors in the IoT and make them available to the terrestrial networks. Coupling SatCom systems with terrestrial cellular networks to integrate new use cases with satellites will provide a powerful new fusion enabling the innovation of services. As SatCom systems can provide an overlay network, the integration of NVF/SDN would enable including network node functions on board satellites to save on physical sites on the ground and open up new chances to improve network resiliency, security and availability. The use of NFV and SDN in SatCom will allow networks to react on demand of the users whenever they are. They also allow the dynamic reconfiguration of the network to give users the perception of infinite capacity of their applications. Satellites would provide a wide coverage area of wireless networks to extend the density of terrestrial cells. They can provide larger cells in heterogeneous arrangements to supply critical and emergency services, take on and keep the network alive in cases of disasters. They also can be able to relieve terrestrial cells of signaling and management functions in a software defined network configuration. Satellites will be integrated to the terrestrial system to improve QoS as well as QoE to end-users. 20 The future SatCom systems will be able to provide intelligent traffic routing among the delivery systems, caching high capacity video to off-load the traffic from the terrestrial networks, and thereby enable saving on valuable terrestrial spectrum. In particular, one of the key drivers of 5G network architecture is the lack of spectrum that would be used for the future wireless infrastructure, so frequency sharing between mobile and satellite systems can deliver major increases in the spectrum provided by both sectors. Leveraging SatCom systems with techniques like SDN, NFV, cognitive and software-defined radio can be built into future systems to allow such frequency sharing. The extension of SDN to satellites would provide an attractive perspective for the SatCom community. By exploiting SDN/NFV satellite equipment will not be vendor-specific; instead they will be open, programmable and reconfigurable platforms. SDN and NFV are expected to offer new cost-effective services, since SatCom operators will be offering the ability to monetize on their network while offering these future/expected services. For example, the emergence of Cloud-RAN would be an enabled for virtualizing SatCom resources (i.e., ground equipment, aerospace access infrastructure), even more the applying NFV and Cloud-RAN to SatCom paves the way towards the full virtualization of satellite head-ends, gateways/hubs and even Satellite terminals, thus entirely transforming SatCom infrastructure, enabling novel services and optimizing resource usage. Network virtualization is considered as the key enabler for the efficient integration of the satellite and terrestrial domains.[6] Via the unified management of the virtualized satellite and terrestrial infrastructures, fully integrated end-to-end network slices can be provided, integrating heterogeneous segments in a seamless and federated way. Additionally, the integration of satellites within the 5G future network will extend the coverage of SatCom systems to support new services such as public transport service, vehicle to vehicle, surveillance with UAVs, high-definition video monitoring, localization and positioning. Moreover, non-geostationary satellites are actually investigated to achieve optimal networking and latency. Intelligent gateways can be designed to improve network resource use by providing hybridization of satellites and ADSL (Asymmetric Digital Subscriber Line) networks. Also, the virtualization can be used to provide black-box (flight data recorder) in the cloud for passenger's aircraft[6]. The role of satellites in the future 5G networks reveals many challenges to support flexible, programmable and secure infrastructure. As satellites will be integrated in 5G broadband networks they should enable extending the coverage of cellular backhaul, while at the same time providing enhanced user-centric QoE, cost-effective user terminals and energy efficiency. SatCom systems should continue to honor guaranteed service delivery to end-users by providing higher throughput and low latency for interactive and immersive services independently from the user location. Additionally, the integration of satellites in 5G networks will introduce new challenges regarding the spectrum sharing. Since mobile terminals will use both terrestrial connectivity as well as satellite connection, mobile receivers should support both kinds of connectivity. Thus, multi-polarized schemes are a key challenge for satellite and context aware multi-user detection.[6] Techniques like SDN, NFV and SDR (i.e., mobile terminals will have 21 modulation and new error-control schemes that will be downloaded from the Internet on the run) are been seen as more challenging aspect of 5G networks so they should be able to provide intelligent orchestration as well as smart antenna beamforming to enable and facilitate frequency sharing between terrestrial and satellite systems.

Delay of Standards / Interoperability

The standards for any technology define what can and can't be done with it. These standards then define how carriers will design and implement their networks. [1]That design then leads to hardware designs (both the radio network and mobile devices) using the standards that are then manufactured, and then ultimately deployed. Though some capital can be expended prior, it is not necessarily in the best interest of the carriers to go out early on their own and presume certain 5G requirements. The more complicated the 5G design, the more important it is to get a solid standard, and that takes time. (Kinney 2015) The amount of interoperability and backwards compatibility will also be a concern. If these networks have to support 4G, 3G, and possible prior generations, it will affect how the system is designed to be able to still communicate with the older technologies, yet remain able to meet the goals of 5G.

The reliance on other communication types, such as WiFi, device-to-device communication and consumer internet service provider access, will cause an entire other set of concerns. Those networks will be required to work together in order to meet the 5G goals.[1] The entire industry of communications, from wireless carriers to wired carriers, and the manufactures of all the devices used by those carriers, will have to come together and agree on how to operate interchangeably across the different networks and establish rules for cost accounting and sharing. These agreements will only add on to the delay created by a late standard.

Conclusion

Many new technologies and a redesign of the wireless network architectures will be able to make the goals of 5G realizable in the near future. The ability and desire for mobile telecommunication companies to implement all of the solutions, however, may require additional solutions to be found . For example, millimeter wave communication technologies may be employed for backhaul traffic . Another idea is to decouple downlink and uplinks. Location information can improve scalability, latency and robustness of 5G . There are also issues related to security given the massive number of connections and quantum techniques might be useful in certain situations especially where one wants to simulate new primitives such as entanglement.

There may also need to be a balance created between these various solutions, depending on the physical environment and capacity density desired. Due to these considerations, portions of 5G requirements may be implemented first, such as additional small cells, before other parts, such as network virtualization, are implemented. A true, fully-implemented-to-the-standard 5G may not be realized for quite some time, and even then, perhaps only in areas that carriers believe they will get the most out of their deployment costs.

Whatever the actual implementation, the technologies are being created and the capability is being developed to be able to implement the 5G requirements in order to meet the mobile communications demands of 2020 and beyond. Early successes in test deployments will further enable modification of systems so that a functional system will be available in the next several years.

References

[1] https://arxiv.org/ftp/arxiv/papers/1705/1705.02902.pdf

[2] https://arxiv.org/pdf/1902.01367.pdf_2020

[3]Goovaerts, D. FCC Unanimously Opens Nearly 11 GHz of Spectrum for 5G. July 14, 2019. https://www.wirelessweek.com/news/2019/07/fcc-unanimously-opens-nearly-11-ghz-spectrum-5g.

[4]https://dspace.networks.imdea.org/bitstream/handle/20.500.12761/771/A_5G_Mobile_Network_Architecture_to_Support_Vertical_Industries_2019_EN.pdf?sequence=1

[5] https://www.mdpi.com/1424-8220/21/16/5578

[6] https://arxiv.org/ftp/arxiv/papers/1506/1506.02876.pdf