5G Reimagined: A North American Perspective (Issue 2)

Alliance for Telecommunications Industry Solutions
February 2017

ATIS–I-0000050.v002

Abstract

The purpose of this white paper is to understand, define, and advance North American requirements for 5G. Deployment scenarios and use cases for 5G networks are analyzed from a North American perspective. These use cases include both traditional and more disruptive service scenarios. The scope of the use cases is not limited to narrowly defined mobile network, and includes interactions with other components.

The white paper identifies unique characteristics of the North American network and regulatory requirements. Although the focus is on the North American market, it is considered in a global context to leverage synergies wherever possible, and to only identify new requirements where necessary.

Foreword

As a leading technology and solutions development organization, the Alliance for Telecommunications Industry Solutions (ATIS) brings together the top global ICT companies to advance the industry's most pressing business priorities. ATIS' nearly 200 member companies are currently working to address the All-IP transition, 5G, network functions virtualization, big data analytics, cloud services, device solutions, emergency services, M2M, cyber security, network evolution, quality of service, billing support, operations, and much more. These priorities follow a fast-track development lifecycle—from design and innovation through standards, specifications, requirements, business use cases, software toolkits, open source solutions, and interoperability testing.

ATIS is accredited by the American National Standards Institute (ANSI). The organization is the North American Organizational Partner for the 3rd Generation Partnership Project (3GPP), a founding Partner of the oneM2M global initiative, a member of and major U.S. contributor to the International Telecommunication Union (ITU), as well as a member of the Inter-American Telecommunication Commission (CITEL). For more information, visit www.atis.org.

Published by

Alliance for Telecommunications Industry Solutions
1200 G Street, NW, Suite 500
Washington, DC 20005

Table of Contents

1 Scope and Purpose 6
1.1 Scope 6
1.2 Business Purpose 6
2 Definitions 7
2.1 5G System 7
2.2 QoS (Quality of Service) 5G System 7
2.3 QoE (Quality of Experience) 7
3 Deployment Scenarios 8
3.1 Introduction 8
3.2 2G 9
3.2.1 Deployment Scenario 9
3.2.2 Business Drivers 9
3.3 LTE Evolution to 5G below 6 GHz 9
3.3.1 Deployment Scenario 9
3.3.2 Business Drivers 9
3.4 LTE Evolution to 5G above 6 GHz 9
3.4.1 Deployment Scenario 9
3.4.2 Business Drivers 10
3.5 Service Migration Requirements 10
3.5.1 Migration and Coexistence Overview 10
3.5.2 Network Combinations (Non-Roaming) 10
3.5.3 North American Migration Trajectories 14
3.5.4 Coexistence Requirements 15
3.5.5 North American Non Roaming Requirements 15

3.5.6 North American Roaming Combinations ... 16
4 Use Cases ... 16
4.1 Network Interoperability Use Cases ... 16
4.1.1 Third Party Wi-Fi Use Cases ... 17
4.1.2 Shared Network Use Cases ... 18
4.1.3 International Roaming Use Cases ... 20
4.1.4 Service Continuity between Devices ... 21
4.2 Privacy Use Case ... 22
4.2.1 Story Highlights ... 22
4.2.2 Business Drivers & Deployment Scenario ... 23
4.3 Policy Based Service Profile Use Cases ... 23
4.3.1 Network Reconfiguration Scenarios ... 23
4.3.2 Low/No Mobility Services ... 24
4.3.3 Automotive ... 25
4.3.4 mHealth & Telemedicine ... 26
4.3.5 Smart Cities ... 27
4.3.6 Mobile Infotainment ... 28
4.3.7 Network Resource Sharing ... 29
4.3.8 IoT Device Focused Services ... 30
4.3.9 User QoE Driven Dynamic Network Use Optimization ... 31
4.4 Regulatory Use Cases ... 31
4.4.1 PSTN Replacement ... 31
4.4.2 PSRA (Public Safety Related Applications) Transition to IP ... 32
4.4.3 Critical Communications ... 33
4.4.4 National Security/Emergency Preparedness (NS/EP) ... 34

4.4.5 Lawful Intercept (including for IoT) 36
4.5 New Business Model Use Cases 36
4.5.1 Local Offloaded Broadband (Roaming Alternative) 36
4.5.2 Flexible MVNO & Carrier Partnering Models 37
4.5.3 Models Trading Access to User Information for Value 39
4.5.4 "All You Can Eat" Content Services (Subscription-Based) 40
4.6 User Experience Use Cases 42
4.6.1 User QoE as a Network Infrastructure Management Tool 42
4.6.2 User QoE as a Marketing & Customer Care Tool 44
4.6.3 Analytics to Predict QoE & Required Network Enhancements 45
4.6.4 Device-Centric Approaches to QoE Use Cases 47
4.6.5 Augmented Reality 47
4.6.6 Virtual Reality 49
5 Network & Architecture Aspects 51
5.1 Network Slicing 51
5.1.1 Discussion 51
5.1.2 Network Slicing Description 51
5.2 Mobility on Demand 52
5.3 Network Aspects of User Experience Use Cases 53
6 Requirements 55
6.1 AR and VR Network Performance Requirements 58
6.1.1 AR and VR Implementation Options 58
6.1.2 Impact of User-Experienced Data Rate 58
6.1.3 Impact of Latency 59
6.2 Mapping of 5G Capabilities to Use Case Requirements 60

7 Security Considerations 63

8 Conclusions/Recommendations 66

8.1 Coherent Evolution 66

8.2 Coverage & Capacity 66

8.3 Enabling IoT, Critical Communications, & Emerging Applications 67

8.4 Implementing 5G for Efficiency & Flexibility 67

Table of Figures

Figure 3.1: 5G Deployment Scenarios 8

Figure 3.2: Possible Scenario for Network Evolution Showing Coexistence of Multiple Options 15

Figure 4.1: Pokemon Go Augmented Reality 48

Figure 4.2: Augmented Reality to Improve Visualization of Medical Situations 48

Figure 4.3: VR Headset and multimode devices 50

Figure 5.1: Network Slices that Cater for Different Use Cases 52

Figure 5.2: Solution Elements for User Experience Use Cases 54

Figure 6.1: Enhancement of Key Capabilities from IMT-Advanced to IMT-2020 56

1 Scope and Purpose

1.1 Scope

The scope of this white paper is to understand, define, and advance North American requirements for 5G. The white paper describes use cases which show, from a North American perspective, possible scenarios for 5G networks. These use cases include both commonly recognized baseline requirements and also more disruptive service examples representing a more challenging conception of aspects of 5G. The scope of the use cases is not limited to just the narrowly defined mobile network. Many of these cases include interactions with other elements, including some not normally standardized, such as content provider applications/networks, operational systems within a carrier network and traffic scheduling and steering algorithms.

Based on the documented use cases, the white paper identifies unique characteristics of the North American network and regulatory requirements. Although the focus is on North American requirements, these are considered in a global context to leverage synergies wherever possible, and to identify new requirements only where necessary.

1.2 Business Purpose

Modern wide area wireless systems were originally introduced as a complementary technology to fixed networks. The initial business model and engineering were designed around the assumption that they would only be used for very occasional high value calls. With each subsequent generation of wireless technology the underlying assumptions changed in terms of traffic volume, service mix, and the role of mobile data. As we enter the era of 5G we see that wireless connectivity is becoming the default mode for many types of users and devices. Wireline subscriptions are declining as users move to a predominantly wireless model. To deal with this emerging service reality, 5G technology must provide a platform for cost-effective provision of high bandwidth, low latency services to the whole range of users.

In 5G we aim to develop a system that is fit for purpose, in a business sense, for the future mix of users and services. Evolution of existing wireless and wireline technologies has created a situation where meeting the full range of modern services requires multiple technology silos to be combined and overlaid in an ad-hoc fashion. As the market drives ever increasing user expectations, this approach is ceasing to be viable. User expectations and increasing service demands will require a coherent approach to technology deployment.

This report explores in detail particular business opportunities and requirements for the 5G network. We aim to provide a unified platform for existing services and also address emerging opportunities including:

- PSTN evolution and wireless PSTN substitution;
- Massive IoT and M2M applications;
- Critical communications; and
- Enhanced mobile broadband.

2 Definitions

For a list of common communications terms and definitions, please visit the *ATIS Telecom Glossary*, which is located at http://www.atis.org/glossary.

2.1 5G System

In previous generations of wireless technology, functionality was primarily provided by network equipment, and devices were simply used to access network functionality. As a result, it was usual to refer to "the network". However, in the case of 5G it is recognized that the boundary between functionality provided by the network and functionality provided by the device is far more fluid. Network Function Virtualization, the role of "apps", and the ubiquitous availability of increasingly sophisticated smart devices will only accelerate that trend. As a result, we generally refer to "5G systems" in this paper to include the combined functionality provided by the traditional network and by the device. However in some cases, the identified functionality is expected to be exclusively in the networks, and in this context, the term "network" is used, rather than "system".

2.2 QoS (Quality of Service) 5G System

Conventional QoS measures transmission parameters such as speed, throughput, latency, jitter, and packet loss. QoS is often used to estimate the service experienced by the user, although in reality it only provides a partial and indirect measure of the user experience.

2.3 QoE (Quality of Experience)

Quality of Experience measures the application and service performance in End User Terms. QoE provides a user/application centric measure of the application's performance that considers factors such as:

- Availability, responsiveness, and infrastructure transparency (i.e., wasted time factors).
- "Intuitiveness": Is it obvious how to learn and use an application?
- Can be sensitive to specific content.
- Value derived vs. time and money spent.
- End User's device(s) performance: this can be a major factor in QoE.
- Ease of use and device portability: these are important factors in QoE.

To accurately determine end-user QoE, one must include performance of application related infrastructure such as OTT servers, Internet backbone and other transport connections, as well as any other elements involved in the overall delivery of the application.

Measuring QoE requires a fundamentally different approach than measuring QoS:

- Different users can measure QoE differently even when the network performance as measured by QoS is identical.
- QoE is dependent on the end user's priority applications and content.
- QoE must be evaluated over target populations.

Accurate measurement of QoE often requires a presence on End-User's device.

3 Deployment Scenarios

3.1 Introduction

Voice calls originate and terminate from many different sources and can be transported over TDM/SS7 (i.e., "circuit switched"), IP/SIP, or a combination thereof. This creates an almost infinite number of calling scenarios involving a myriad of technology combinations.

A consideration of possible 5G deployment scenarios is a first step towards identifying service continuity requirements. This clause proposes 5G deployment scenarios as the first step in this analysis. These scenarios are similar to use cases, but they are centered on the deployment issues rather than user experience issues. Therefore they are described in a separate deployment scenarios section that provides context for all of the 5G use cases.

As shown in the figure below, it is anticipated that initial deployments of 5G will likely be primarily in urban areas with high capacity demand. Over time coverage will spread to include areas with lower user density.

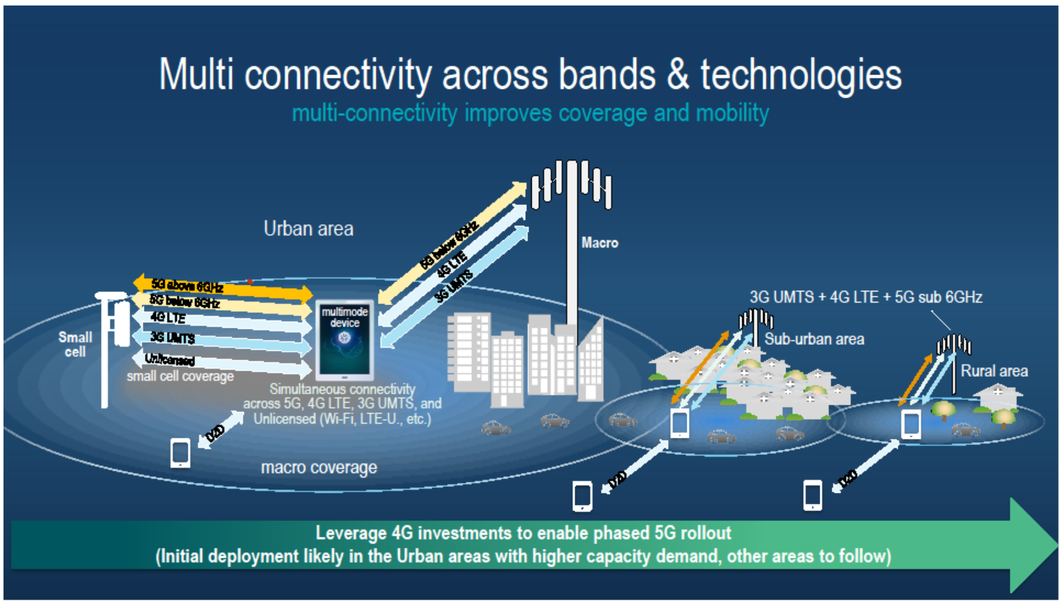


Figure 3.1: 5G Deployment Scenarios

3.2 2G

3.2.1 Deployment Scenario

2G networks are still widely deployed. However by modern standards 2G technology has limitations in security and capacity that are difficult to avoid. It is therefore anticipated that many operators will want to retire 2G technology and reuse the spectrum for more modern systems by the time 5G reaches deployment. In some cases, 2G technology may continue to exist to support existing niche applications, but these would not be expected to interoperate with 5G systems or be supported by 5G devices.

In this scenario interoperability and handover between 5G and 2G is not required. To avoid known security vulnerabilities, 5G devices may not be permitted to attach to 2G networks.

3.2.2 Business Drivers

Replacing installed 2G equipment will require a capital investment though in some cases this will occur naturally as network equipment is replaced for other reasons. Offset against this cost is reduced operating costs due to the greater efficiency of newer technologies and simplification of the network configuration by removing one generation of technology. Deployment of newer technologies with higher capacity will also allow more services to be offered to users on new technologies.

3.3 LTE Evolution to 5G below 6 GHz

3.3.1 Deployment Scenario

Deployment of 3G UMTS wireless systems is still growing in many countries and the technology is expected to continue in service well beyond 2020. In parts of the world, UMTS could continue to be the primary technology used for worldwide roaming. 3G UMTS has a number of advantages over earlier technology generations, including spectrum efficiency, capacity, support for packet switched data, and enhanced security. As a result, the incentive to replace UMTS technology could be small, and it is expected that the technology will co-exist alongside 5G for some time.

3.3.2 Business Drivers

Replacing installed 3G UMTS systems will require capital investment to replace equipment that in many cases is adequate for users, in particular for roaming users who may use this technology in their home networks. UMTS systems support most of the basic services offered by later generations of wireless technology, though not necessarily with the same level of performance. If there are no compelling triggers to upgrade (e.g., severe spectrum shortages), many service providers will opt to continue operating 3G systems alongside 4G/LTE and 5G systems. Although it is expected that 3G systems will coexist alongside 5G systems in some regions, and that 3G will be important for roaming in some markets, the need to handoff sessions between 5G and 3G is for further study.

3.4 LTE Evolution to 5G above 6 GHz

3.4.1 Deployment Scenario

Above 6 GHz 5G will be used to provide very high capacity in areas of dense users and where there is a short distance between users and the network antennas. This is required due to the propagation characteristics of high frequency radio. The services and service continuity requirements for 5G operating above 6 GHz are for further study.

3.4.2 Business Drivers

5G deployment above 6 GHz will be driven by use cases where the capacity requirements cannot be satisfied with the available spectrum below 6 GHz. The propagation characteristics in these frequency ranges will, at least initially, make this technology suitable for targeted deployments, rather than a general deployment over wide geographic areas.

3.5 Service Migration Requirements

3.5.1 Migration and Coexistence Overview

Migration from legacy to 5G includes both non-roaming and roaming aspects. The non-roaming aspect describes North American requirements for migration and co-existence within a single operator's network. The timeline for the various transitions are controllable by the operator to minimize disruptions. The roaming aspect covers requirements for operating in external networks at different levels of maturity in the migration to 5G. In many cases, the timelines for these scenarios are much longer.

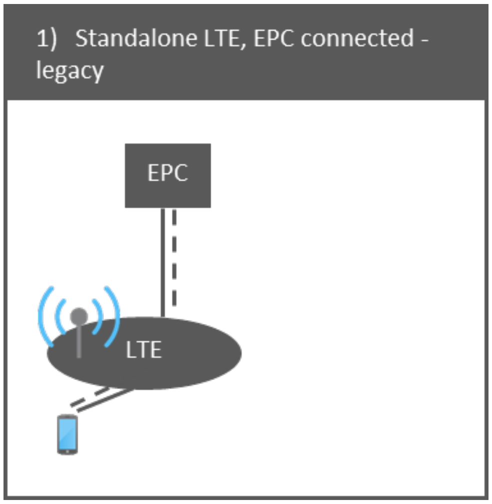
3.5.2 Network Combinations (Non-Roaming)

Even within a single operator's network, it is unreasonable to expect that the network will migrate to 5G as a flash cut. 5G consists of a radio component (New Radio [NR]) and a network component (Next Generation Core [NGC])

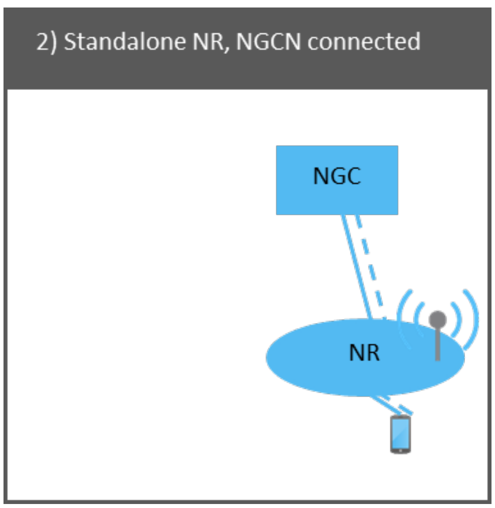
3GPP has broken the possible combinations into various scenarios called *options*. These capture various combinations of the radio and core network as well as how the signaling is routed between the components. Not all of these options make sense.

The diagrams below show the various options. In these diagrams:

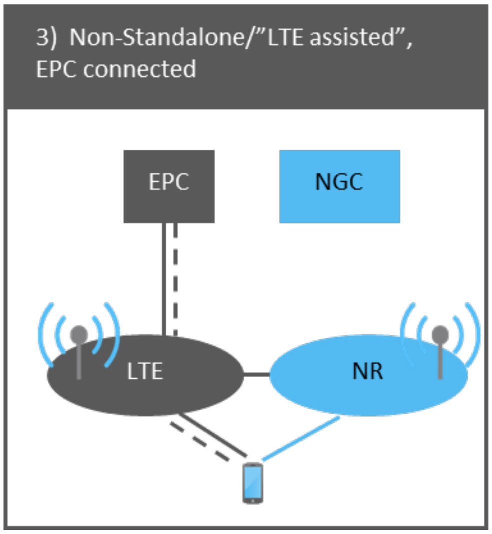
- "EPC" is Enhanced Packet Core i.e. the LTE core network.
- "LTE" is the LTE RAN.
- NextGen Core/NGC is the proposed new core network for 5G.
- New Radio/RN is the proposed new 5G RAN.
- Solid lines show user plane traffic.
- Dashed lines show signaling traffic.

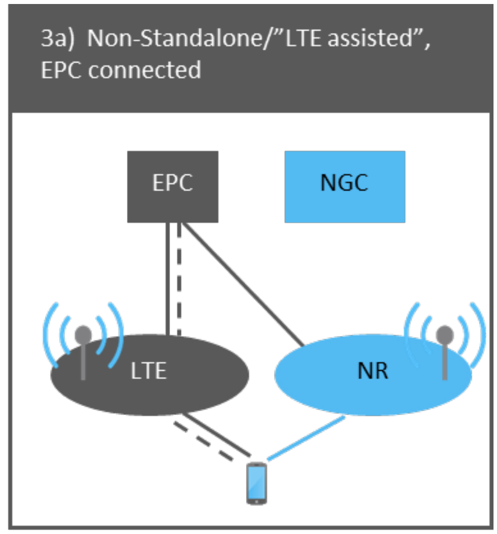


Option 1 represents a legacy network. This is the starting point.

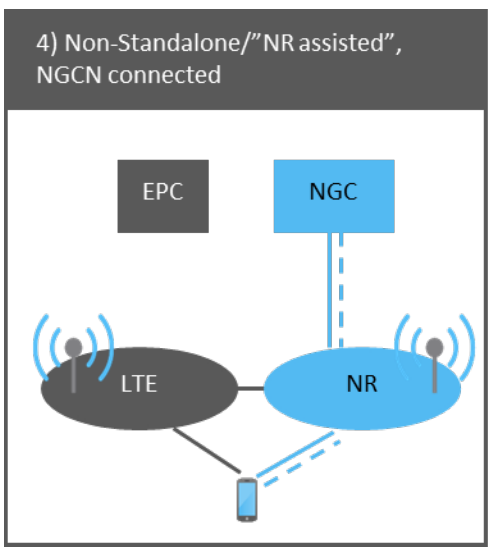


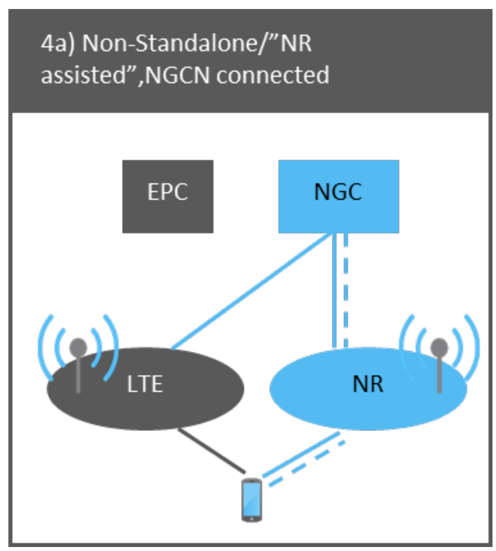
Option 2 represents a greenfield new network. It has no LTE components. It may also be a network once LTE is completely phased out; however, the phase out of LTE is likely to take some time.



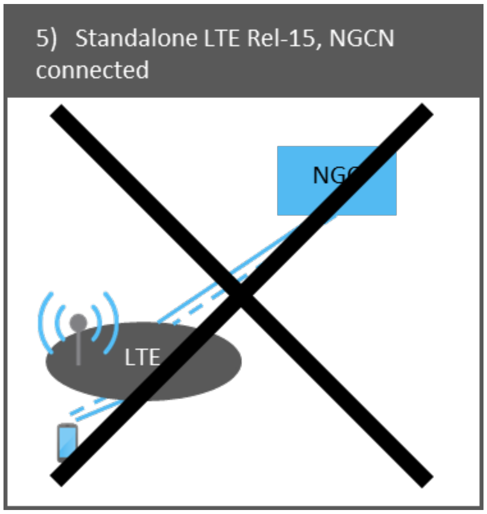


Option 3 represents a network using only the EPC core. It has both LTE and NR radio access. Control signaling is routed through the LTE RAN. Options 3 and 3a differ based on whether the user plane data is sent to NR directly or via the LTE RAN. This is a likely first step in the migration toward 5G.

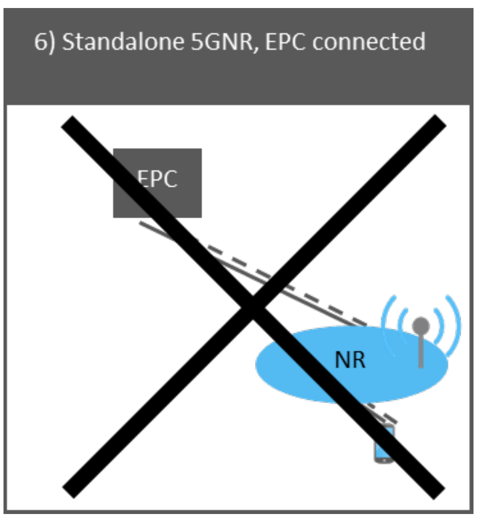




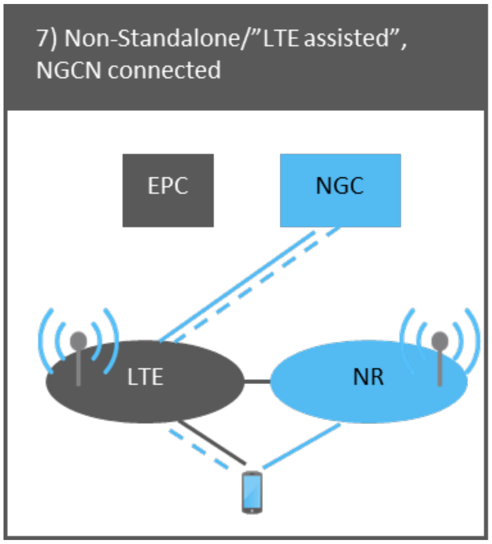
Option 4 represents a network using on the Next Generation core (NGC). Control signaling is routed through the NR RAN. Options 4 and 4a differ based on whether the user plane data is sent to LTE directly or via the NR RAN.

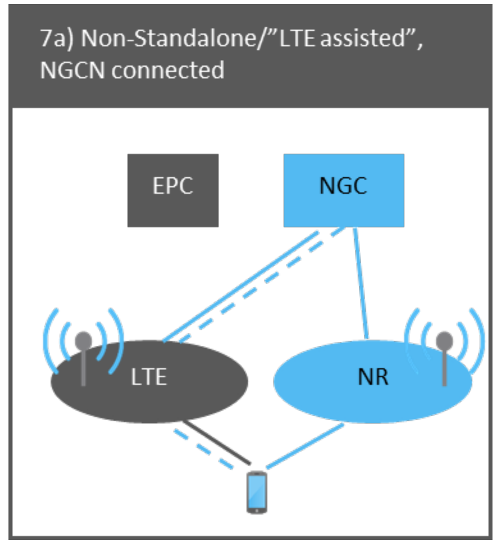


Option 5 represents a network that has transitioned toward the NGC, but continues to use LTE access. In this case LTE is an evolved LTE RAN that understands new signaling. This option seems unlikely as most of the benefit of 5G will derive from migrating towards a new radio.

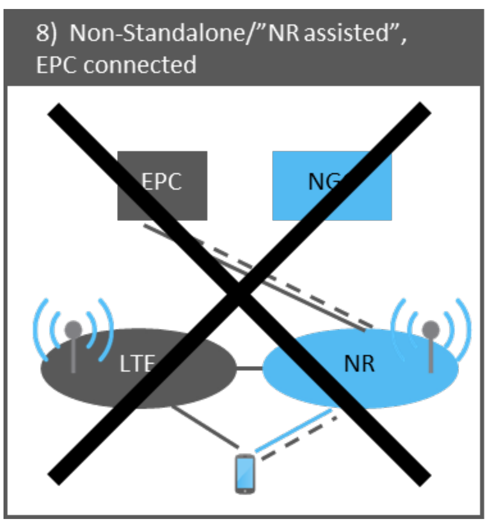


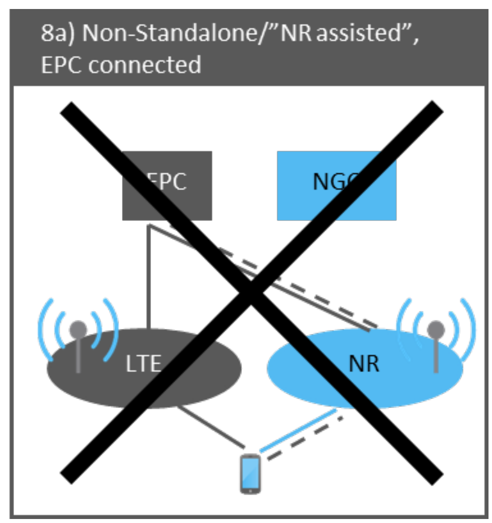
Option 6 represents a network that has migrated to the NR core, but retains EPC as the core network. This uses original EPC signaling between the EPC and NR. This is considered a less likely scenario than Option 3 because any operator having EPC is likely to already have a LTE network. In the time it takes to phase out LTE, it is likely that the core network will migrate to NGC, making this scenario unlikely.





Option 7 represents a network using on the Next Generation core (NGC) and a mixture of LTE and NR radio. Next generation signaling is used; however, the signaling is routed via the LTE RAN. This is because the LTE network is more fully built out and is thus more reliable for handling signaling. Options 7 and 7a differ based on whether the user plane data is sent to NR directly or via the LTE RAN.





Option 8 represents a network using only the EPC core. It has both LTE and NR radio access. Control signaling is routed through the LTE RAN. Option 8 is considered unlikely since an operator's LTE coverage is likely (at least initially) to be larger than their NR coverage. For this reason, it makes no sense to route signaling via the NR RAN. Options 8 and 8a differ based on whether the user plane data is sent to LTE directly or via the NR RAN.

3.5.3 North American Migration Trajectories

In North America, one possible trajectory for migration to 5G is:

Step 1. Option 1 (Legacy)
Step 2. Option 3 (NR introduced, but using EPC)
Step 3. Option 7 (NGC introduced). However; LTE is the dominant network and still handles control signaling
Step 4. Option 4 (NR becomes the predominant network). Control signaling is now via NR.
Step 5. Option 2 (LTE phased out)

It is understood that operators can stop anywhere along this path, and the interval between each step may be highly variable. Operators may also choose other trajectories depending on their circumstances and business objectives.

3.5.4 Coexistence Requirements

The options from 3GPP describe combinations of a single UE in terms of connecting to a RAN and a core network. However, a commercial network will likely have a large population of UEs and deployment situations. Thus, multiple options will need to be supported simultaneously within an operator's network depending on UE capabilities and what is deployed within a given area. This describes which of these options will need to be supported.

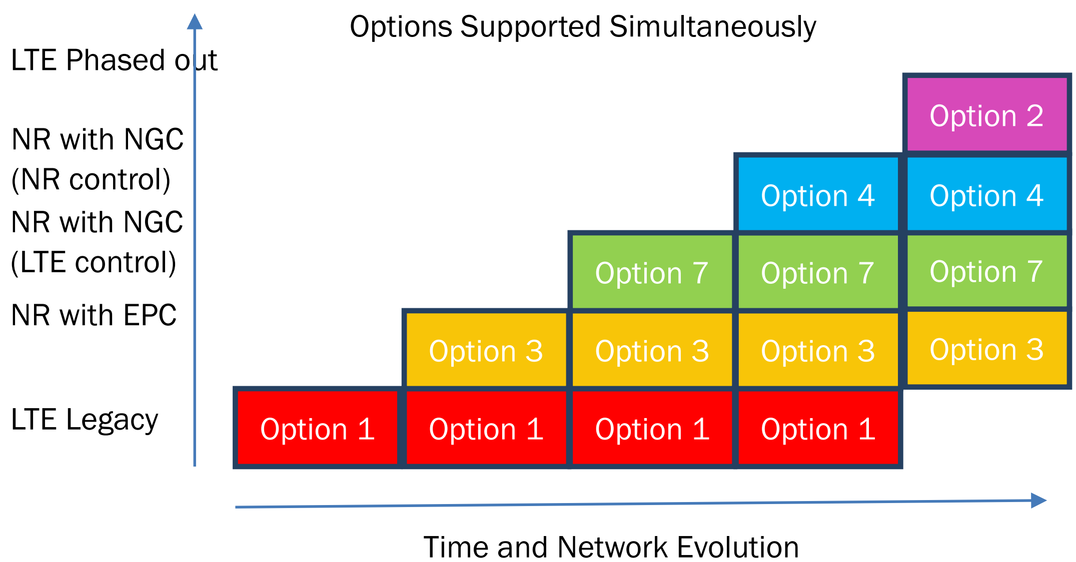


Figure 3.2: Possible Scenario for Network Evolution Showing Coexistence of Multiple Options

The above diagram shows that the networks will need to simultaneously support many of these combinations for a long period of time.

3.5.5 North American Non Roaming Requirements

To allow the above migration it is proposed to introduce the following requirements:

- 5G UEs will be able to support both EPC and NGC signaling.
- 5G UEs that cannot support NGC signaling will always initially attach via LTE.

Based on the UE capabilities (or lack thereof), the RAN will steer the UE to the appropriate core network

3.5.6 North American Roaming Combinations

The following key scenarios are considered:

- A LTE UE roaming into a 5G network.
- A 5G UE roaming into a LTE network.

A LTE UE roaming into a 5G network – this scenario is supported if:

- The serving network has LTE radio, and
- The serving network has an EPC core (option 1 or 3) OR the LTE UE has been upgraded to support NGC NAS signaling (option 4).

A 5G UE roaming into a LTE network – This scenario is supported if:

- the 5G UE supports the LTE RAN, and
- the 5G UE supports the LTE NAS (option 1 or 3) OR the serving LTE network has been upgraded to a NGC core (option 5). This scenario is unlikely.

4 Use Cases

Requirements for 5G are derived from use cases that consider how the 5G system will provide services to the end user. The use cases are grouped into broad categories that better illustrate the requirements that derive from these use cases.

4.1 Network Interoperability Use Cases

A number of use cases relate to scenarios where the user selects from a number of network options, depending on service, performance, cost, or user preferences. When the service transfers from one network configuration to another, one of the important considerations is the degree of network interoperability. The following use cases outline representative scenarios where network interoperability is an important consideration.

4.1.1 Third Party Wi-Fi Use Cases

5G will include advanced support for carrier integrated Wi-Fi and other technologies in unlicensed spectrum. However it is also required that 5G coexist with third-party Wi-Fi services that have no carrier integration. Users may, in some cases, prefer third-party Wi-Fi services for reasons of cost, convenience, feelings of control, or to provide special access to local or enterprise services. The following use cases show some scenarios.

4.1.1.1 Residential Home Network (Non-Integrated)

4.1.1.1.1 Story Highlights

Adam has a home Wi-Fi network which is not integrated with his mobile service provider. When Adam enters his home network area, his mobile automatically selects his home network for IP based services that are supported while maintaining a connection to the mobile network for value added services not provided on a non-integrated Wi-Fi network.

In this use case, services provided by the user's home Wi-Fi network are not billable by the mobile operator. Services provided by the mobile network while also connected to Wi-Fi are billable as normal.

4.1.1.1.2 Business Drivers & Deployment Scenario

Users will continue to want to make efficient use of their home network resources in the 5G timeframe. Independently deployed Wi-Fi networks should be supported, sometimes in preference to the mobile network.

However, not all 5G services may be available on simple Wi-Fi connections. Therefore in order to deliver the full range of services, both Wi-Fi and mobile connections should be simultaneously supported.

4.1.1.2 Enterprise Network (Non-Integrated)

4.1.1.2.1 Story Highlights

Bettina works in an office building with a Wi-Fi network which is not integrated with her mobile service provider. When Bettina enters her enterprise network area her tablet, which is 5G equipped, automatically selects her enterprise network for IP based services that are supported while maintaining a connection to the mobile network for value added services not provided on a non-integrated Wi-Fi network. Maintaining a connection to the enterprise network is important because some private enterprise services are only available to users connected over their enterprise network.

Bettina is able to access both service provider hosted services and enterprise hosted services while she is connected to both networks.

In this use case, services provided by the enterprise Wi-Fi network are not billable by the mobile operator. Services provided by the mobile network while also connected to Wi-Fi are billable as normal.

4.1.1.2.2 Business Drivers & Deployment Scenario

As with the residential home network case, independently deployed Enterprise networks should be supported.

4.1.2 Shared Network Use Cases

The coverage and capacity goals of 5G require deployment of extensive radio infrastructure. Sharing of infrastructure is one approach that may be adopted to reduce costs. Shared infrastructure is also relevant where physical, environmental, or other constraints make the building of multiple infrastructure networks difficult or impossible. The following use cases show scenarios where different models of network sharing may apply.

4.1.2.1 Integrated Enterprise Network in Multi-Tenant Building

4.1.2.1.1 Story Highlights

Chris works in a multi-tenanted office building. The building's landlord has deployed a "neutral host" 5G wireless system in the building that allows any 5G service provider to offer services. Chris' employer has made an agreement with a network operator, ACME NET©, to provide enterprise services both at the office and in the public environment.

When Chris is outside his office, his mobile is connected to ACME NET's© public infrastructure. When Chris enters his office his phone automatically connects to the private neutral host system but with services still provided by ACME NET©.

In this use case all services provided to Chris are billable by ACME NET©. However it is likely that the arrangement between ACME NET© and Chris' enterprise is such that special rates apply. This could, for example, include unlimited data provided Chris is at his office location. There will also be a billing relationship between ACME NET© and the neutral host. This relationship may include a fixed fee for the provision of connectivity as well as charges for services used. 5G systems should efficiently support the billing, auditing, and payment settlement of such arrangements.

4.1.2.1.2 Business Drivers

This use case illustrates several possible business opportunities in 5G systems:

- The use of a single provider by an enterprise for both public mobile services and a hosted service within the enterprise's office environment.
- The use of "neutral hosts" to allow many operators to serve managed spaces such as shared office buildings and shopping centers with a common infrastructure.
- The use of the 5G technology umbrella to provide services that today are typically provided by Wi-Fi.

4.1.2.1.3 Deployment Scenarios

This use case illustrates one of several scenarios for how enterprise services might be offered in the 5G timeframe. 5G should be capable of supporting the density of users and bandwidth required to provision enterprise-class services in an office environment. This would mean that where there is abundant 5G capacity it may not be necessary to deploy a traditional Wi-Fi overlay.

For difficult to cover managed spaces such as offices, tower blocks, shopping centers, and sports arenas, some kind of neutral host may become a popular deployment scenario. The neutral host would deploy the necessary infrastructure and wholesale access services to mobile operators. This would enable mobile operators to reach users while avoiding the costs and difficulties associated with supporting multiple operator-specific infrastructures.

Neutral host systems will require access to suitable spectrum. This may involve operating in unlicensed spectrum or it may be that service providers who offer services on neutral hosts permit the neutral host to operate in frequencies that they have licensed.

4.1.2.2 IoT Not Spots

4.1.2.2.1 Story Highlights

Fenix Inc. is a provider of monitoring equipment for storage tanks at gas stations. The company deploys sensors that regularly report the level of gas in the tanks and monitor the tanks for leaks and other safety hazards. Fenix Inc. has made an agreement with ACME NET© to provide connectivity to their gas tank sensors.

Some rural gas stations equipped by Fenix Inc. are in locations where ACME NET© does not have mobile coverage (not spots). However coverage is available from a community supported 5G system.

By default the Fenix Inc. sensors connect to the ACME NET© coverage. However in "not spots" they will connect to any suitable access network.

In this arrangement ACME NET© will need to collect sufficient information to bill Fenix Inc. and ensure that their devices conform to any agreed terms (e.g., a maximum number of data updates in a given period). ACME NET© and third party networks used to provide coverage will need a commercial relationship. This relationship may include a fixed fee for the provision of connectivity as well as charges for services used. 5G systems should efficiently support the billing, auditing, and payment settlement of such arrangements.

4.1.2.2.2 Business Drivers & Deployment Scenario

For both mobile operators and IoT users it is advantageous to make a single national or regional agreement to provide IoT connectivity. Recognizing that a single operator is unlikely to cover every location, operators will often form national roaming agreements with other parties to improve their effective coverage for IoT users.

4.1.2.3 Human User Not Spot

Discussion: In Europe there is growing political pressure for operators to cooperate in order to solve the problem of rural "not spots". Operators are starting to respond to this by taking initiative to improve coverage which may also include use of shared infrastructure and national roaming. As part of establishing the North American requirements for 5G, recognizing that many areas in North America have even greater challenges related to large areas with very low population density, it is appropriate to address this in the North American context. The deployment scenarios and business drivers will be similar to the IoT Not Spots use case above.

4.1.3 International Roaming Use Cases

5G will provide users with global services based on the ability to roam and access local infrastructure in visited countries. The ability to balance user control and convenience associated with roaming is an important factor for users. Another important aspect is the consistency of services users experience when they are not in their home country.

4.1.3.1 Casual Consumer International Roaming

4.1.3.1.1 Story Highlights

Denise is taking a vacation abroad. While on vacation she wishes to continue to access social media sites and multimedia entertainment but wishes to manage roaming charges. Denise sets the preferences on her mobile to reflect this. While travelling her mobile only automatically connects to services which are compatible with her preferences. This may include connecting to services such as:

- Free public Wi-Fi services.
- Denise's operator's preferred roaming partner where usage is counted within Denise's normal monthly allowance.

In some locations Denise is unable to connect because there is no suitable provider available. Denise is happy to restrict her usage to manage roaming charges.

Sufficient information will need to be collected to bill Denise in this scenario. Information will also be required to support inter-operator billing. In addition to conventional inter-operator billing other payment scenarios are possible, for example payments could be automatically deducted from Denise's electronic wallet by the local providers.

4.1.3.1.2 Business Drivers & Deployment Scenario

Many mobile operators have preferred roaming partners, perhaps within the same group of companies. Users wishing to manage roaming charges should be steered towards these services.

Popular leisure destinations are often well provided with free Internet services both at tourist businesses (e.g., hotels and restaurants) and sometimes in public spaces as part of municipal initiatives. Users wishing to manage roaming charges will want to make the best use of these facilities when travelling.

4.1.3.2 Best Connectivity International Roaming

4.1.3.2.1 Story Highlights

Eric is a director of a large company. When travelling he wants to be constantly connected even if this incurs additional charges. Eric sets his preferences on his mobile to be "best connectivity" roaming. When travelling, his mobile automatically connects to any suitable connectivity service though it may prefer lower cost services where they are suitable and available. For example:

- While on a plane Eric's mobile will automatically connect to the in-flight Wi-Fi if his operator has a roaming agreement.
- While in a foreign country Eric's mobile will automatically connect to any roaming partner of his operator, though it will prefer roaming partners specified by his operator.

Sufficient information will need to be collected to bill Eric in this scenario. Information will also be required to support inter-operator billing. This scenario illustrates that 5G operators may provide services via a very wide range of third parties both at home and abroad. The need to manage these relationships in a simple, cost effective and secure manner will be an important feature of 5G.

4.1.3.2.2 Business Drivers & Deployment Scenario

For users demanding the best connectivity, continuous access to services even on buses, trains, and planes, this approach will become the expectation. Operators will make multiple roaming agreements to satisfy this demand. In some scenarios there may be a choice of partners to serve a user, depending on user preferences. Both technical and business aspects may feed in to the decision on which partner is used.

4.1.4 Service Continuity between Devices

One goal of 5G is to provide a unified experience for users across multiple devices. From an operator's point of view 5G also aims to unify multiple access types into a single core network. As well as requirements for interoperability for one device operating on different access technologies, this also introduces requirements for service continuity between devices.

5G systems should provide the user with the ability to transfer existing service sessions from one device to another device. This transfer should take place with a minimum perceptible interruption to the user.

4.1.4.1 Audio Streaming Use Case

A user is listening to an audio streaming service on their in-home audio system. When the user leaves for work she wishes to continue listening to the same service in the car. This may involve several steps – for example initially the service may transfer to the user's mobile phone while she walks to her car. Once she starts her car the service may transfer from the mobile on to the car's audio system.

A mechanism is required to enable the user to indicate when she wants the transfer to take place. There are several possible ways this can be achieved including:

- Manual request from the user interface of one or the other device. This would require some mechanism to allow the user to identify and associate both devices.
- Automatic behavior based on pre-defined rules such as location or reachability.
- By placing the relevant devices in close proximity to associate them and using a special user interface event to initiate the transfer.

For the audio streaming service a short, user perceptible interruption may be acceptable.

4.1.4.2 Conversational Service Use Case

A user is listening to a conference call while driving to work. Once he reaches his office he wishes to continue the conference call on his speaker phone. In addition he would now like to connect his desktop PC to the shared presentation screen for the conference.

The transfer could be initiated using mechanisms similar to those described above.

For the conversational service a short, user perceptible interruption may be acceptable.

4.2 Privacy Use Case

4.2.1 Story Highlights

User selects a network for the purpose of accessing public Internet services. How can the user be assured that:

- Her communication with an intended party remains private and confidential, specifically;
  - Data exchanged remains private/confidential and
  - The identities of those communicating remain private.
- The identities are mutually authenticated. Assured that they are communicating with the entity/person they believe they are communicating with (in both directions).

4.2.2 Business Drivers & Deployment Scenario

The 5G system must intrinsically provide these assurances so that one does not have to rely on the client/host for basic security needs. Privacy would be assured by the service provider.

This then opens the question as to Lawful Intercept (LI) and other matters of national security. This is for further study.

4.3 Policy Based Service Profile Use Cases

It has been suggested that one feature in 5G is that the profile for services may be defined using an approach based on context awareness and network policy rather than by the deployment of different service networks. So, for example, normal and critical communications uses may be served by the same service nodes with the service type provided being selected based on the context (e.g., device type) and user's policy attributes.

The following use- cases illustrate various policy-based network selection scenarios.

4.3.1 Network Reconfiguration Scenarios

4.3.1.1 Story Highlights

An operator wishes to streamline its response to various scenarios that can occur. These situations often occur within a given geographic area. The types of scenarios covered include both predictable and non-predictable events.

- Predictable: Concerts, parades and races, sporting events, rush hour, black Friday, etc.
- Non-predictable: Hurricanes, major traffic accidents, floods, earthquakes, etc.

Typical actions that might occur in such a scenario include:

- The reallocation of virtual resources within the network for purposes such as increasing capacity, increasing redundancy, introducing different types of VNFs, etc.
- Alter the network topology to handle failures, increase capacity, etc.
- Altering the service mix to prioritize certain types of services, restrict access, and introduce specialized services.

4.3.1.2 Business Drivers & Deployment Scenario

Reconfiguration of a network to respond to new scenarios can be a time consuming process with today’s technology. Automating much of this process would save both time and reduce operational costs.

4.3.2 Low/No Mobility Services

4.3.2.1 Story Highlights

5G is expected to see an explosion of IoT devices as well as a variety of consumer and enterprise devices. These devices will require diverse mobility requirements driven by specific applications. A variety of IoT applications leveraging cellular infrastructure can be envisioned; opportunities exist from power meters used in Smart Grid to Public Warning Systems utilizing wirelessly connected earthquake/tsunami detection sensors. All such applications can and are beginning to be deployed even on today's cellular networks. However IoT applications are predicted to grow at a much faster pace than perhaps what existing networks and cellular technologies can optimally handle. To support billions of IoT devices, a wireless network infrastructure is needed which is not only highly scalable in terms of its capacity, but can also optimally handle differing service needs of various IoT verticals. Examples of differing service needs include different requirements on mobility, as well as latency, network reliability, and resiliency. This diverse set of requirements may require re-architecting key components of the cellular network, for example to support mobility-on-demand whereby mobility is only provided to those devices and services that need it.

4.3.2.2 Business Drivers

- The presence of a very large number of IoT devices on the cellular network could require radically new technology and solutions that the current 4G networks may not support.
- Low or no mobility devices will also extend beyond the IoT space into consumer and enterprise devices.
- To support mass scale IoT deployments, scalability is required on the device as well as the infrastructure side.
- 5G network must be designed to support connectivity to a massively larger number of low-cost, low-power, simple devices and device types in the "Internet of Things".
- A key requirement of IoT communication is low power consumption while still maintaining a high degree of reliability and coverage.

4.3.2.3 Deployment Model

- The network architecture must allow for different core network instances, each optimized to support a different type of device/traffic profile. Such a feature would not only enable scalability but will also help reduce cost.
- In 5G architecture, traffic profile parameters may include: mobility modes, latency, bandwidth, reliability, etc. NFV orchestration can be used to create these packet core instances.
- The 5G network architecture and air interface should be designed in a manner that will allow machine type devices to co-exist on the same network as other high data rate devices, while still allowing for maximum battery life and power efficiency.
- The network should be designed in a manner to dynamically trade-off between spectral efficiency and power/energy efficiency depending on the device type and the traffic profile.
- The concept of peer-to-peer connection to some other higher-powered gateway device could be supported.
- Other considerations include:

- o Appropriate support for mobility based on device and service.
- o Extreme low-mobility/short duration and infrequent transactions (e.g., IoT sensor scenario).
- o Connectionless access.
- o Upstream poll-only downstream access.
- o Flexible definition of profiles, i.e., device can create a new one on the fly.
- o Consider shifting frequency utilization/RAN access platform selection as part of the profile (i.e., profile completely independent of access/spectrum available, net chooses the best available access).
- o Lower-latency/power and other items that are currently under examination in the industry; may want to consider a more consistent connection to the Internet.

4.3.3 Automotive

4.3.3.1 Story Highlights

Growing populations, more cars, and limited highway capacity result in severe traffic congestion. Advanced Driver Assistance Systems (ADAS) and Autonomous Vehicles will bring a number of benefits including better safety, fewer collisions, less congestion, better fuel economy, and higher productivity. 5G will enable real-time collection and processing of massive amounts of data from vehicles and sensors deployed throughout the city to help streamline traffic flow. Additionally, 5G wireless technologies supporting high-speed low-latency vehicle-to-vehicle and vehicle-to-infrastructure communications will enable cooperative vehicles and pre-crash sensing and mitigation.

4.3.3.2 Business Drivers

Traffic congestion is a major issue in most urban and even rural areas, and leads to productivity loss, environmental pollution, wasted fuel, and degradation of quality of life. Even more tragically, more than 30,000 people die each year in the United States alone from automobile accidents. Collisions lead to injuries and the associated high cost of health care (>2.5M emergency room visits and 1M hospital days), lost productivity (>$33B lifetime from just 2012 accidents), and property damage. Traffic accidents also result in time and productivity loss due to traffic congestion. The best way to keep people safe and reduce the senseless waste of time and money is to prevent crashes from happening in the first place.

4.3.3.3 Deployment Scenario

5G will enable real-time collection and processing of massive amounts of data from vehicles, drivers, pedestrians, and wireless traffic sensors deployed throughout the city to help streamline traffic flow. Traffic signals can be adapted to road usage, road use fees/tolls can be adjusted to throttle traffic in congestion zones, and public transportation can be directed to where it is needed most. Adaptive vehicle navigation systems can use the data to minimize transit time and reduce driver cost in fuel, time, and tolls.

5G wireless technologies supporting high-speed low-latency vehicle-to-vehicle and vehicle-to-infrastructure communications will enable ADAS and fully Autonomous Vehicles. Cooperative vehicles operating safely as a platoon on highways will improve highway capacity, reduce congestion, eliminate the occurrence of driver error, and achieve better fuel economy. Pre-crash sensing will enable vehicles to sense imminent collisions and exchange relevant data among vehicles involved. ADAS systems will allow vehicles and drivers to take counter-measures to eliminate or at least mitigate the impact of collisions. Pre-crash sensing and platooning requires highly reliable and extremely low latency (<20ms) vehicle-to-vehicle communications. In addition, vehicle to pedestrian communication can provide a valuable mechanism to further reduce traffic fatalities. 5G should support the scalability and capacity to enable massive amounts of data exchange among a large number of vehicles, e.g., in a congested highway segment, without significant impacts to communication performance and reliability.

Automotive communications services also require a highly secure, reliable, and rugged wireless broadband network.

4.3.4 mHealth & Telemedicine

4.3.4.1 Story Highlights

Medical information can be efficiently exchanged and patient care cost effectively delivered using 5G technology and various types of user equipment. Doctors can extend care to patients at home, hospitals can provide care to patients in remote areas, and first responders can communicate with health care professionals in emergencies. Elderly or immobile patient consultations using high-definition video conferencing instead of office visits, transmission of ultra-high resolution MRI images, patient portals to provide and receive information, remote monitoring of vital signs, and 24/7 real time urgent care nursing call centers are just a few examples of how 5G enabled telemedicine will reduce health care costs, minimize response times, and improve the overall level of care.

4.3.4.2 Business Drivers

Today, the U.S. spends 15% of its GDP on healthcare, and reducing or containing the cost of healthcare is a top national priority. Telemedicine has been shown to reduce the cost of healthcare and increase efficiency through better management of chronic diseases, better utilization of health professional staffing, reduced travel times, and fewer or shorter hospital stays.

Telemedicine can provide service to millions of patients from a centralized location with a smaller staff, can address local provider shortages, and can help protect highly trained professionals from dangerous situations. Instead of staffing and maintaining expensive remote facilities, 5G communication will allow physicians and health facilities to expand their reach without staffing and maintaining inefficient and expensive remote facilities.

Some healthcare services delivered via 5G telemedicine will be as good as those provided by traditional in-person consultations, and will reduce lost productivity from travel time, and long patient queues at health facilities. In some cases, such as mental health and critical care, e.g., triage and stroke assessment, telemedicine can deliver immediate intervention resulting in superior outcomes, patient satisfaction, and reduced unnecessary emergency room visits.

4.3.4.3 Deployment Scenario

5G will enable future medical applications through significant improvements to wireless network data throughput, capacity, latency, security, and reliability. Ultra-reliable 5G wireless networks will link patients with doctors, and urban hospitals with clinics, rural health centers, and mobile first responders. With 5G connectivity, medical records containing high resolution medical images and video can be made available to physicians and medical professionals anytime, anywhere. Improvements in the quality of low-cost mobile cameras, displays and application software, portable diagnostic tools, and automatic medical telemetry via wearable sensors, will further increase the demand for bandwidth on 5G systems.

Medical sensors and diagnostic tools can include life critical applications for cardiac, pulmonary, or fetal monitoring in a hospital or home care. Ultra-reliable wireless connections will be used to communicate directly between the patient and the monitoring center in real-time, and enable rapid intervention by a health care professional. However, applications can also include informational or recreational monitors to modify lifestyles or improve athletic performance. The application aware 5G wireless systems will need to differentiate between life critical and informational data traffic, and determine QoS based on the application, location, context, user or device subscription, and user preferences.

Life critical, personal, and confidential health services require a highly secure, reliable, and rugged wireless broadband network.

4.3.5 Smart Cities

4.3.5.1 Story Highlights

Today's cities depend on a wide array of critical infrastructure to function properly; electricity, water, sewer, gas, etc. Critical infrastructure monitoring is an expensive undertaking, often requiring service levels achievable only by dedicated wire-line connectivity. For instance, in order to detect a fault in high-voltage transmission lines, and be able to take corrective action to prevent cascading failures, the required communication latency is beyond what current wireless networks can achieve. Similarly, structural monitoring requires the provisioning of a large number of low-data-rate battery-powered wireless sensors, and today's wireless networks are not optimized to support this deployment model, both in terms of battery life and cost efficiency.

Also, with the massive migration of the world's population toward urban environments, there is increasing demand on cities to modernize their infrastructure and services. From water and power management to buildings and transportation, city planners will rely on new scalable, interconnected services that don't require cities to overhaul private and public infrastructure and are built for the future.

5G will enable cities around the world to build long term connectivity strategies that help improve livability and sustainability.

4.3.5.2 Business Drivers

Malfunction or damage to critical infrastructure, such as electricity, natural gas, or water, could result in a huge financial impact, quality of living degradation, and even loss of life. The power blackout in Northeastern United States in 2003 was an example of how infrastructure failure can bring an entire region and its economy to a halt. Other examples of such disruption include bridge and building structural failure leading to collapse, water and sewer system malfunctioning, etc. It is therefore important to monitor the “health” of critical infrastructures reliably and cost effectively.

4.3.5.3 Deployment Model

Massive numbers of 5G connected sensors and actuators will collect and process data to monitor critical parameters and optimize performance based on current conditions. Sensors will also enable service providers to detect when hidden pipes and cables need repair, and when unauthorized access takes place.

5G will be designed to support reliable low-latency communications among densely deployed devices that are subject to power constraints and wide-ranging data rate requirements.

4.3.6 Mobile Infotainment

4.3.6.1 Story Highlights

Video and audio streaming, video calls, social networking, and multimedia messaging are just some of the popular communications and entertainment applications used on today’s wireless networks. Additionally, new applications will emerge, such as real-time multi-user gaming, virtual/augmented reality, 3D multi-site telepresence, high-resolution (4K & 8K) video streaming, picture and video sharing, etc. These applications will require significant increases in data rate, capacity, and ultra-low communication latency that is not supported by today’s wireless networks. 5G will be the solution that enables the continued evolution of communications and entertainment applications.

4.3.6.2 Business Case

Cisco[1] and Ericsson[2] predict that U.S. mobile data demand will grow by roughly a factor of 6 over the next five years to 2020. With the 3G and 4G enabled proliferation of smartphones and other mobile data devices, we have all come to expect Internet connectivity everywhere—at home, at work, in our automobiles, and wherever we go. The increasing adoption of mobile video and other data intensive applications is placing significant strain on today’s 3G and 4G networks. At event venues, stadiums, and other traffic hotspots with high concentrations of active wireless users, today’s networks can be severely congested leading to an unacceptable user experience. The voracious consumer appetite for anytime and anywhere information and entertainment will ensure that the exponential growth in mobile data demand will continue. How to support the rapid anticipated growth in mobile data traffic in densely populated areas will be an important driver of 5G requirements.

---

[1] Cisco Visual Networking Index (VNI), May 2015: < http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html >.

[2] Ericsson Mobility Report, June 2015: < http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf >.

4.3.6.3 Deployment Model

More access nodes, improved spectral efficiency, and more spectrum must be leveraged in tandem to satisfy the rapidly growing demand for bandwidth. Improved utilization of current frequency bands <6GHz and new technologies such as higher order MIMO, advanced carrier aggregation, closer integration of Cellular and Wi-Fi, NSLO, CoMP, and direct device to device communication will be implemented in future versions of 4G. But to meet demand, 5G will need to deliver even more effective use of licensed and unlicensed spectrum, ultra dense network deployments using ever higher frequency spectrum, and efficient allocation of wireless resources based on application and context. 5G will also expand on the techniques developed in 4G for new network architectures, secure protocols, smart antennas, and much more.

4.3.7 Network Resource Sharing

4.3.7.1 Story Highlights

Network sharing business models involve a relationship between the mobile service provider and the network operator, and between network operators, in which their respectively owned physical network infrastructures are tightly coupled through a sharing arrangement. The system should provide the capability for various network sharing schemes to maximize the overall synergies of network sharing agreements and to enable flexible business models and commercial relationships that potentially may change rapidly.

4.3.7.2 Business Drivers

- Network sharing is emerging as a disruptive mechanism to control network deployment costs.
- Although several components of the mobile network can be shared, RAN (Radio Access Network) sharing is believed to have the most impact.
- Dynamic sharing of resources among carriers/MVNOs (requires mechanism to reserve/allocate and release resource in real-time).
- Applies to all parts of Infrastructure (RAN, core, billing and sub-management infrastructure, security, authentication, other levels).
- Reservation/use/release of spectrum.
- Enable “service enabling plane” sharing.
- Public or land-lord owned infrastructure sharing (e.g., “small cell” access in building or public places).
- Virtual broker platform for venue owners and operators.
- Broker of access to other key infrastructure – poles, buildings, types of venues.

4.3.7.3 Deployment Model

- Provide sufficient flexibility to accommodate the capacity needs of dynamically hosted operators on a real-time basis (e.g., for capacity brokering architecture, where network resources are provided dynamically depending upon bids offered).
- Technical capabilities shall include spectrum sharing or reuse, enhanced mobility techniques, and enhanced controls for access network, access point, access node, and spectrum selection at an operator policy level.

4.3.8 IoT Device Focused Services

4.3.8.1 Story Highlights

There is a great deal of interest and early work around enabling smart, connected devices and the creation of an Internet of Things (IoT). This use case is about enabling growth through the following:

1. Reducing the cost of connectivity of Smart Devices to support massive deployments.
2. Enabling carrier networks to provide connectivity for both cellular and non-cellular IoT devices.
3. Enabling the efficient use of spectrum for peer to peer devices.

This use case builds upon the ideas proposed in the “Lo/no Mobility” and “Local Offloaded Broadband” use cases to provide greatly improved cost, lower complexity, and improved efficiency for IoT applications.

4.3. 8.2 Business Drivers

- Accelerate uptake by enabling a more efficient ecosystem.
- Expand the addressable IoT market for carriers.
- Enable efficient P2P solutions.

4.3. 8.3 Deployment Model

- Dedicated spectrum to support a more efficient ecosystem could be realized through a nationwide IoT Spectrum Block (along with appropriate service profiles).
- Expanding the addressable market through simpler protocols would be deployed by essentially allowing non-cellular devices to be integrated into a carrier’s network trivially and automatically, with minimal to no effort. This could be at the device level or through interfaces to the devices cloud applications (an implementation plan has not been defined yet). As such, this might require some enablement of device ID sharing between IoT device providers and carriers. For example, ZigBee devices ideally would be plug and play into a carrier core network via Wi-Fi to a carrier’s router, or ZigBee devices known to their own cloud application could be interfaced via Internet into the carrier’s network trivially.
- P2P communications using dynamically managed spectrum would be deployed with increments to existing D2D protocols that would enable the allocation and then release of spectrum in real-time for short bursts of data transfer.

4.3.9 User QoE Driven Dynamic Network Use Optimization

4.3.9.1 Story Highlights

As 5G networks start to deploy it will be a fact that multiple RATs and heterogeneous architectures will be the norm. Traffic balancing has become more important to operators as spectrum has diversified into non-contiguous blocks that may not offer ideal propagation or bandwidth for the service that the user has requested. Additionally, the IoT and other machine to machine communications will become more prevalent and will need to be handled in ways that do not interrupt critical real time user needs. To accomplish this more advanced traffic steering, it is suggested that network analytics could be used to anticipate the need to reserve network attributes for the application or service that may be about to be run. If there are limits to what services the network can offer the user or device, the network could also report those limitations to the user or device. This would be useful in allowing the user or device to make another application/service choice or to reduce frustration with a perceived lack of service.

4.3.9.2 Business Drivers

1. Better spectrum utilization.
2. Reduction in network congestion.
3. Improved customer expectations.
4. Reduced customer complaints.
5. Better UE human factors design.

4.3.9.3 Deployment Model

There are two ways that analytics and predictive use behaviors could be used. A network could use either one by itself or both. The network can provide, based on current usage, what level of service, e.g., QoE, QoS, that it can provide to a user or user equipment. This is then used by the UE to allow certain services to be requested. In the meantime, until the service could be provided by the network, the UE would store data for later forwarding and would notify the user of this delay and suggest other applications that could be used. For example, at a sporting event, data bandwidth could be in short supply. The network would state to the UE that only text based services are currently available. If a user were to try and upload a picture the user would be told that the upload would take place later and that text is the only available service at this time. This approach can be further elaborated by expanding conventional SON framework to an end-to-end level that also incorporates user traffic demand and behavior in combined cognitive radio/machine learning.

4.4 Regulatory Use Cases

It is important to note that "regulatory" does not imply that the functionality described here is required in all use cases. Some regulatory requirements only apply to selected services and are not required if the service provider does not offer the service.

4.4.1 PSTN Replacement

4.4.1.1 Story Highlights

As more and more users "cut the cord" and move to wireless only service, the remaining customers face increasing pressure to also make the transition. Although the latest technology can offer exciting new services, some customers need, or highly value, existing attributes of the PSTN. This includes high reliability, as well as support for "legacy services", such as PSRA, that are highly valued by segments of the population. Ensuring that 5G technology can meet critical requirements of the PSTN will be important to ensuring that 5G is seen as an acceptable replacement technology.

4.4.1.2 Business Drivers

PSTN infrastructure is rapidly aging with much of the equipment reaching end of life, or already manufacturer discontinued. In many cases the original manufactures are no longer in business, making support and spares increasingly problematic. In addition, a declining customer base supporting an increasing cost structure will at some point become unsustainable. It is important to ensure that 5G can meet key PSTN service characteristics to allow it to be seen as a credible technology for PSTN replacement at the appropriate time.

4.4.1.3 Deployment Scenario

If standard 5G infrastructure satisfies PSTN service requirements, it will be possible to migrate PSTN lines to 5G over time based on local conditions.

4.4.2 PSRA (Public Safety Related Applications) Transition to IP

4.4.2.1 Story Highlights

The public safety sector covers the broad set of Fire, EMS, Police and Emergency Operations Centers (FEPE) related applications currently provisioned over legacy copper infrastructure. Given the diversity of the applications, the provisioning methods include circuit switched voice (POTS), 4 wire leased lines, analog fax and modem lines, and data services such as DDS. While integration with Next Generation E9-1-1 and emergency services requirements is important in this area, this use case is not focused on specific E9-1-1 requirements, as this is being addressed in other industry initiatives.

As existing networks migrate to wireless solutions, and as wireless migrates to 5G, it is important to ensure that 5G supports PSRA requirements.

4.4.2.2 Business Drivers & Deployment Scenario

A common thread across the public safety sector is the need to achieve a high level of reliability and maintain a consistent continuity of operations with any technology migration. Public safety entities are facing the same infrastructure challenges as network operators, i.e., many current products (in this case customer equipment) are being discontinued and/or replaced in the marketplace with new products that offer higher capacity, new interfaces built to new standards, and a wider set of media options. Public safety entities are presented with various alternatives to migrate to IP-enables solutions, but must assess the full complexity of each solution as it relates to their specific operation.

It is this challenge that has formed the basis for further assessment and information sharing covering IP migration paths for public safety related applications.

4.4.3 Critical Communications

Next generation critical communications depend heavily on advances in telecommunications. Concepts that are considered as essential in the future include:

- Video cameras.
- Hazmat sensors.
- Medical sensors.
- Connected vehicles/roadways.
- Augmented reality.
- Security.

As a case study for public safety, let's consider the next generation first responder as shown in Fig. 4.1.

Fig. 4 - 1: The Next Generation First Responder[3]

[3] Source: DHS S&T APEX Program

Wearable technology will require fast, reliable, secure, and high bandwidth communication between all first responders. The head-mounted display should be able to show in real-time an augmented reality of the area along with the location of all the other responders, as well as the location and status of the threats.

Sensors require ultra-reliable and low latency communication. As an example, the sensors on a first responder must be able to detect the environmental threats and send the relevant data, along with the responder's health information, to a central command center where the data will be analyzed and potential alerts are sent to all the responders. This will require low latency, high reliability, secure, and high throughput communication.

In addition to public safety, 5G must enable advanced communication between individuals and authorities. As an example, one should be able to send high definition video of an incident to a Public Safety Answering Point (PSAP) and receive the relevant instruction or information.

Finally, as sensors are deployed in roads, bridges, railroad tracks, buildings, etc. to enhance safety of the public, it is important to remember that some IoT devices will require ultra-reliable and secure communication with potentially some priority level.

4.4.4 National Security/Emergency Preparedness (NS/EP)

NOTE: Though NS/EP is categorized as Authority to Authority critical communications, it encompasses both Authority to Authority and Authority to Individual priority communications (i.e., an authorized NS/EP user can initiate a priority communication to any called party).

The 5G system will have to support the NS/EP priority communication services, including the ones that are being deployed for Next Generation Network (NGN), and support backward compatibility/interworking with the existing legacy NS/EP services (i.e., Government Emergency Telecommunications Service (GETS) and Wireless Priority Service (WPS)). In addition, as new 5G devices, applications, and network capabilities are defined and become commercially available, it shall be possible to use these 5G devices, applications, and network capabilities to support end-to-end priority NS/EP communication services, as appropriate.

General NS/EP objectives and use cases for 5G include, but are not limited to:

- 5G support of backward compatibility/interworking with existing legacy NS/EP communication services, i.e., GETS and WPS.
- 5G support of Next Generation Network Priority Service (NGN-PS) being deployed for NGN. NGN-PS is the evolution of legacy GETS and WPS, through technology insertion, to achieve service continuity in a packet-switched NGN and leverages the NGN to offer new features, including priority voice, video, and data services.
- Use of new commercially available 5G devices, applications, and network capabilities to support end-to-end priority NS/EP communication services, as appropriate. For example, it is possible that there may be a need to use an IoT application for NS/EP. This means that it should be possible to use the commercially available IoT device and application to provide a corresponding NS/EP use of the IoT application in the public 5G system. Similar concepts would apply to new 5G user and enterprise applications, such as tactical

multimedia services.

- Support of more efficient ways to provide priority treatment and Quality of Service (QoS) for NS/EP communications in a shared 5G public network environment. For example, 5G should explore more efficient methods to prioritize traffic independent of QoS characteristics (e.g., latency). Specifically, in a shared public 5G network environment, it will be necessary to enforce relative priority and precedence treatment among multiple critical applications, based on national and service provider policy.
- High availability of NS/EP communications in the 5G environment. This will depend on robustness, reliability, and resiliency of the public 5G communication infrastructure to support NS/EP communications. It is desirable that capabilities for robustness, reliability, and resilience be supported in the 5G system that will in turn provide the needed high availability for the supported NS/EP application services.
- Use of 5G capabilities such as Network Function Virtualization (NFV) and slicing to provide efficient support for the deployment and operations of NS/EP application services. For example, this may include use of NFV capabilities for dynamic control of network resources for NS/EP, including support of service recovery operations and restoration priority requirements.
- Integrity, confidentiality/privacy, and availability protection of NS/EP communications and NS/EP users in the 5G environment.

4.4.4.1 NGN-PS Voice Service

4.4.4.1.1 Story Highlights

A hurricane has hit a major metropolitan city. A FEMA (authority subscribing to NGN-PS Voice Service) government official tries to make a regular cellular voice call to local officials in the metropolitan city to assess the situation but is unable to contact them. He receives a fast busy signal, finds that the destination phone number does not ring, or hears a message that the call cannot be completed. The networks are congested and the probability of completing a normal call is reduced.

The FEMA government official tries the call again by using a WPS-subscribed cellular phone. The official dials the local officials in the metropolitan city using the WPS access code. The 5G system recognizes the call as a WPS call and provides the subscribed priority functions. If this call is handed off to another service provider, the service provider will recognize the WPS call and provide the priority treatment for the call. The FEMA official is able to contact the local officials, assess the situation, and start provisions for FEMA support functions.

4.4.4.2 NGN-PS Data Service

4.4.4.2.1 Story Highlights

A hurricane has hit a major metropolitan city. A FEMA (authority subscribing to NGN-PS) government official arrives at the metropolitan city to assess the situation. The FEMA official uses his LTE smart phone to access a FEMA web server so that he can download assessment procedures. Due to network congestion, network connectivity to the server is timing out. The FEMA official then invokes the NGN-PS Data service to achieve data network connectivity, acceptable QoS, and throughput for subsequent downloads from the server. Invocation is done using a browser on his WPS subscribed smartphone.

4.4.4.3 NGN-PS Video Service

4.4.4.3.1 Story Highlights

A hurricane has hit a major metropolitan city. A FEMA (authority subscribing to NGN-PS) government official arrives at the metropolitan city to assess the situation. The FEMA official needs to have a video conference with the FEMA team back at the headquarters to discuss the situation. He uses his smart phone to place a normal video call to the conference bridge number. The video call is not connecting due to network congestion. The FEMA official then invokes the NGN-PS Video service and obtains a priority connection to the video conference bridge. The audio and video quality is good and the conference video call proceeds as planned.

4.4.5 Lawful Intercept (including for IoT)

Existing LI requirements will continue to apply in the context of 5G, but it is possible that new devices and services for IoT could impose additional requirements to support LI. Although IoT devices are very different than, for example, smart phones, they both share key attributes including:

- Devices have a network address in the context of any given session
- Devices can be contacted
- Communications can be established with the device

Therefore, at this time it is not expected that any additional technical requirements will be needed for 5G systems to support LI for IoT devices.

4.5 New Business Model Use Cases

4.5.1 Local Offloaded Broadband (Roaming Alternative)

4.5.1.1 Story Highlights

The scenario here is to create a class of services that do not require connection-maintaining mobility and can deal with being offloaded to the Internet at the RAN level. This approach has the potential to significantly lower the cost of a high percentage of mobile broadband traffic while simultaneously improving latency and user experience. If this class of service were made available at a substantially lower cost, it is likely that many of today’s mobility applications would be revised to take advantage of this service profile.

4.5.1.2 Business Drivers

- 70% of users are not mobile, and IoT growth will likely increase the level of non-mobile users.
- Non-licensed competitors will likely focus on this segment of the market for expansion.
- Enables limited mobility solutions like Wi-Fi to be used to reduce costs and increase bandwidth.
- May ease initial use of millimeter wavelength spectrum to enhance coverage and bandwidth.
- End-user ability to absorb increased costs for increased levels of mobile broadband consumption is limited.
- Competitive and regulator pressure to create flat-rate roaming scenarios.

- Service profile well suited for flat-rate and zero rated billing models.
- Enables or simplifies a broad range of emerging MVNO and flat-rate roaming scenarios (see related use cases).

4.5.1.3 Deployment Model

- Locally administered core functions (e.g., connection, IP address assignment, RAN QoS, perhaps limited mobility).
- Usage/billing data is the only item that is transferred to centralized "administrative mobile core". The end user's device could be used as a source for usage data.
- This service profile should be independent of RAN type and spectrum selection but can take advantage of many existing unlicensed RAN architectures including Wi-Fi and others.
- User device and associated applications will negotiate for QoE and other service parameters such as session persistence, constant IP addresses, etc. with network.
- Compatible with licensed spectrum assisted models being discussed today (e.g., LTE-U).

4.5.2 Flexible MVNO & Carrier Partnering Models

4.5.2.1 Story Highlights

There is a great deal of interest in business models which enable more efficient use of infrastructure, spectrum, and other resources which make up wireless infrastructure. There is also considerable interest in cooperative business models involving carriers, content providers, and MVNOs. This use case addresses these opportunities via a variety of breakthrough approaches to billing and customer care coupled with a flexible and lighter weight approach to infrastructure sharing, control, and management. The use case includes the need to accommodate alternative billing models for "flat-rate", "zero-rated", mutual "all you can eat", and standard billing models as well as a customer experience management and "care" platform which unifies provider owned content and networks along with shared wireless infrastructure. This use case could be applied to video, voice, broadband data, or IoT wireless services. Many of the customer, service, and business scenarios enabled by this use case will also make use of our Local Offloaded Broadband, Low/No Mobility Services, and Network Resource Sharing Use Cases.

4.5.2.2 Business Drivers

- Vertical market segments such as content providers and IoT applications service providers would be able to focus on their customers and service experiences without having to address the complexity of building and managing complex wireless infrastructures.
- Carriers will be able to share each other's infrastructure as well as access to 3rd party networks (e.g., Wi-Fi in buildings) in a way that provides a seamless end-user experience to their customers.
- Content providers are looking for mechanisms to utilize wireless infrastructure to deliver their content and services in a manner which can guarantee a high-quality user experience. At the same time, these providers often wish to maintain a primary, one-stop customer care and billing relationship with their customers.

- Wireline-centric and other fixed broadband and unlicensed wireless providers are looking to create a seamless MVNO-style relationship with wireless operators to provide wireless and mobile extensions to their service sets.
- Significant improvements in end-user experience are possible using the “SmartPipe” services and seamless sharing of 3rd party networks that would be enabled by this use case.
- A much higher level of dynamic sharing of wireless infrastructure and spectrum would be enabled which in turn would significantly lower costs while simultaneously improving the utilization of precious resources such as wireless spectrum.
- The potential to deliver a much improved end-user experience for content (e.g., entertainment video) and other services would be enabled by this use case.

4.5.2.3 Deployment Model

- An MVNO, content provider, or a service provider engaged in network sharing would realize a highly integrated service delivery capability via a light weight platform providing customer care, experience management, and billing services. This platform would interoperate with a partner network that would provide access and possibly mobile core capabilities. Such a lightweight billing and customer care platform might be realized via a cloud/NFV-based architecture. This approach would enable a content provider, a carrier engaged in network sharing, or an MVNO to deliver wireless services without building a conventional packet core or RAN network to cover some or any of their service footprint.
- An operator should be able to provide mobile service by building any combination of the required elements – billing/customer care virtual core, packet core, and RAN – and sharing the rest via appropriate business relationships with other operators.
- The virtual billing/end-user care component should be architected to be “light weight” and cloud based to facilitate integration with content service platforms which already exist on the Internet (e.g., Netflix®, Hulu™.)
- There will need to be some changes in existing service models to accommodate some of the flexible roaming arrangements outlined above (e.g., carriers hosting roaming would not be required to apply home network services).
- Interfaces between the virtual core and the packet core/RAN should enable a variety of service profiles and billing models to be applied to realize a “SmartPipe” capability for content and other performance sensitive services.
- As in other use cases, the network should decouple the service profile from the specifics of RAN type and spectrum chosen to deliver a service.
- The business models outlined here would be realized using a combination of licensed, unlicensed, and license assisted access methods which meet the requirements of the associated service profiles.
- This use case could be greatly facilitated by user device-driven billing models which are secure and reliable. This area needs to be explored in more detail.
- This use case would likely be utilized in conjunction with our Local Offloaded Broadband, Low/No Mobility, and Network Resource Sharing Use Cases to efficiently realize a range of content and other services.

4.5.3 Models Trading Access to User Information for Value

4.5.3.1 Story Highlights

Increasingly, models which trade reductions or free content or other Internet-based services for access to information about end-users service and content usage are becoming common. This is especially true for consumer-centric services. Carriers need to be able to realize this type of business model to be competitive in this environment. While these types of trade-offs are applicable to a wide range of content delivery and service models, we will focus on video content delivery in this use case as:

- Business models which trade access to user information are already commonly used as part of many video services (e.g., YouTube™, many cable and satellite channels).
- Video is a service that is often provided by carriers (in part or as a whole).
- Today's video service models usually integrate anchor content from the primary content provider along with additional content from other services which are not part of the primary content lineup distributed by the provider.

The idea is to create a flexible platform for end-users to determine what parts of their usage pattern they wish to share with the content providers involved. Such a platform would also allow end-users to see what they might be able to receive in terms of rate reduced or "free" services in exchange. The platform would also be able to allow trade-offs to be made on a content category or provider specific basis. For example, a given end-user might be more open to sharing information with his or her primary content and broadband wireless provider but less willing to sharing information with a bundled 3rd party video service which they rarely use anyway. The types of information that end-users might allow (or not) to be shared could include:

- Their location;
- The services which they use and when they use them;
- The specific content they are viewing (this would require co-operation between the carrier and the content provider if not the same); and/or
- Information about the Quality of Experience (QoE) that they have during the viewing of specific content (this would involve a prompt to collect the information at the end of a viewing session).

Perhaps all but certainly the last two items above would require interaction with the user and the device-based client associated with the content service being used. It is also likely that an "opt-in" application would need to be provided on the end-user's device.

4.5.3.2 Business Drivers

- Understanding end-user usage patterns and deriving value from the same is required to be competitive as a full stream content provider in today's markets.
- Carriers are in a strong position as anchor video content providers to create an alternative ecosystem for brokering the use of end-user information if they can provide a platform with proper opt-in and protection capabilities.
- These models provide a means to derive value from 3rd party content providers that are increasingly important to marketing a competitive content lineup.

4.5.3.3 Deployment Model

- An application and associated set of APIs for use in content consumption clients would be required on the popular end-user devices.
- A suitably robust billing and ratings system would be needed to enable per-user, per-application customization of charging.
- A back-end system to reconcile credits and to provide controlled access to end-user usage patterns would be needed.

4.5.4 "All You Can Eat" Content Services (Subscription-Based)

4.5.4.1 Story Highlights

Peak data rates expected by users double every 12 to 18 months, downloads by the top users exceed terabytes per month, and an increasing percentage of users (currently in excess of 20%) regularly exceeding 100 gigabytes of data use per month. On the lower end of the customer base, the volume of data used by the lower half has not changed in several years.

This tends to show that unlimited data services are likely to appeal to an increasing portion of customers. 5G networks are following the demand for continually increasing capacity and connectivity, but at a cost, which makes it important to consider specific business cases associated with the highest network capacity drivers.

4.5.4.2 Business Drivers

5G networks will be used to connect an increasing number of customers to more and more applications. We could start to create a collaborative environment between network providers and the application providers. Network providers may offer multiple new 5G plans and options. Applications will continue to evolve and will likely include a future variant of the following:

- Streaming media.
- File transfer (transfer from cloud and game downloads).
- Messaging and collaboration.

4.5.4.3 Deployment Model

Deployment models can be differentiated in accordance to various parameters.

- Zero rating:
    - Assumption:
        - The regulatory environment evolves to allow for carrier/content provider relationships that could drive preferential content to specific QoS control.
    - Concept:
        - A wireless consumer establishes a relationship with a content provider whereby that content provider's customers receive service experience based upon a contractual agreement between the carrier and the content provider.
        - The wireless consumer would not experience any usage metering while utilizing a "Zero Rated" service.
        - The wireless carrier would receive bulk compensation from the content provider.
    - Normal service levels are available, while contractually negotiated services may be architected differently.
    - Customers may be prompted to be informed of a special application rate, and optionally given the choice to subscribe to it.
- Application specific network resources (and QoS):
    - Changes in QoS may occur according to applications, and/or zero rated services.
    - Opportunities for high speed local (RAN) caching of content.
    - Force handoff to wireless of higher order video compression transport streams, content aware compression.
    - Statistical previewing/real-time sharing of RAN queue information.
- Machine learning:
    - 5G networks may have some advanced knowledge of content demand and its location. Applications can use that knowledge to improve QoE, anticipate resource needs, and for network optimization.
    - Network may push some content closer to the edge before a request is made – to local RAN or even UE storage.
    - Application providers can anticipate next video/asset based on content choice on a given page – network providers can anticipate where that asset is likely to be requested. The combination of the two is a powerful tool, and can be used for great QoE improvements, network efficiencies, and bringing assets closer to the edge – even start trickling to the device.
- Mobile and fixed resources:
    - Content distribution across multiple simultaneous radio interfaces:
        - Licensed and Unlicensed assisted, bursting to local UE queued cache.
        - Traffic Sensitive RAN Offload: optimize network and stream content to a less loaded adjacent site.
        - Upstream operations may continue to the normally desired serving site (interference control).

4.6 User Experience Use Cases

4.6.1 User QoE as a Network Infrastructure Management Tool

4.6.1.1 Story Highlights

While the network controlled and class-based Quality of Service (QoS) concepts of the 3GPP Evolved Packet Core (EPC) are based on the 3GPP policy and charging control (PCC) framework, Quality of Experience (QoE) research has shown complementing QoS mechanisms with more user-centric approaches is needed to truly meet end-user requirements and expectations[4]. ITU-T[5] defines QoE as "the overall acceptability of an application or service, as perceived subjectively by the end-user". Wikipedia defines QoE as:

"A measure of a customer's experiences with a service. QoE looks at a service offering from the standpoint of the customer or end user, and asks, "What mix services, and support, do you think will provide you with the perception that you are being provided with the experience you desired and/or expected?" It then asks, "Is this what the service provider has actually provided?" If not, "What changes need to be made to enhance your total experience?" In short, QoE provides an assessment of human expectations, feelings, perceptions, cognition and satisfaction with respect to a particular product, service or application."

In a communications system, QoE can be viewed as attempting to cover the entire ecosystem, including technology, business, and human behavior aspects[6]. Network operations looks at network performance attributes and quality of service parameters, while the business end looks at attributes such as average revenue per user and customer churn, and finally behavioral scientists look at end user satisfaction.

QoE may be used to help answer "what changes need to be made to enhance your total experience?". This is the essence behind incorporating QoE as a network management tool. As we move to 5G, network slicing, and virtualization of the network, the network will be more and more context aware of the end user/device needs. This context awareness may also include QoE parameters associated with the device or service in order to provide an optimized experience to the end user, and may be based on access location, time of day, or type of service/application. QoE may also have the potential "to become the guiding paradigm for managing quality in the cloud"[7].

QoE requirements and resulting technical requirements from applications differ in various usage dimensions which include the degree of multimedia intensity,

---

[4] "Survey and Challenges of QoE Management Issues in Wireless Networks", Sabina Baraković and Lea Skorin-Kapov, Journal of Computer Networks and Communications Volume 2013 (2013).

[5] ITU-T: "Definition of Quality of Experience (QoE)", International Telecommunication Union, Liaison Statement, Ref.: TD 109rev2 (PLEN/12), Jan. 2007.

[6] "Quality of Experience in Communications Ecosystem", Kalevi Kilkki, Nokia Siemens Networks, Journal of Universal Computer Science, Volume 14/Issue 5 (2008).

[7] "Challenges of QoE management for cloud applications", Hobfeld, T.; Univ. of Wurzburg, Wurzburg, Germany; Schatz, R.; Varela, M.; Timmerer, C., IEEE Communications Magazine, Volume:50 , Issue: 4 (2012).

interactivity, primary usage domain, and service complexity.

4.6.1.2 Business Drivers

- Improve customer/end-user experience by providing an optimized service experience.
- Identify optimal network resources for a particular end user or service:
  - Optimum QoE has to be achieved while constraining the application to behave as resource-efficiently as possible in order to minimize operational costs.
  - Allows for better dimensioning of network and radio resources.
- Identify performance and other network issues.
- QoE based data used to make recovery decisions during outages or other network degraded conditions.

4.6.1.3 Deployment Model

Examples from the literature:

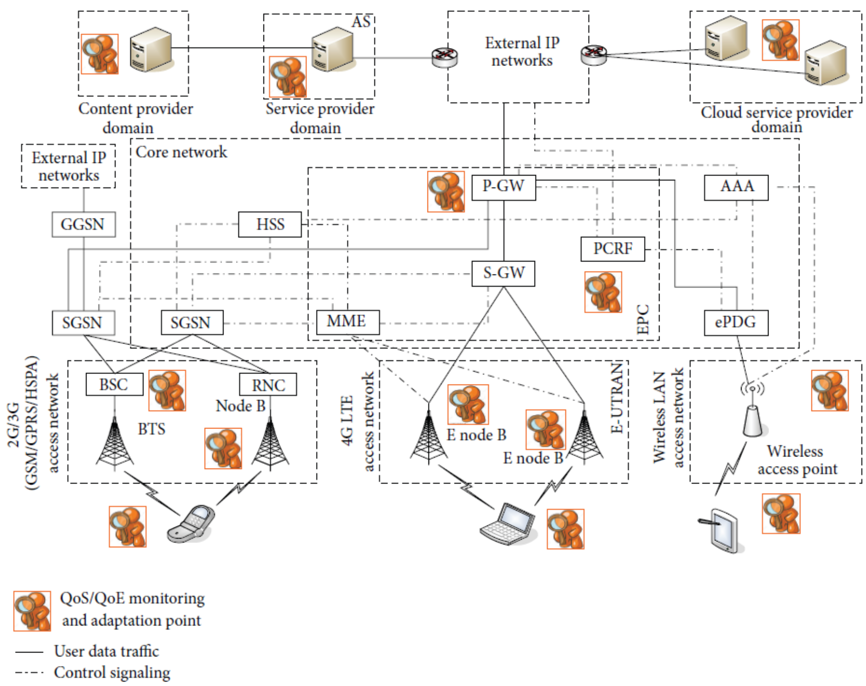


FIGURE 2: QoE/QoS monitoring and adaptation points in a wireless network environment (see Abbreviations).

Figure 2, 5, and 6[8] illustrate possible QoE/QoS monitoring and adaptation points in a 3GPP EPC network. Similar concepts can be designed into the 5G network.

[8] "Survey and Challenges of QoE Management Issues in Wireless Networks", Sabina Baraković and Lea Skorin-Kapov, Journal of Computer Networks and Communications Volume 2013 (2013).

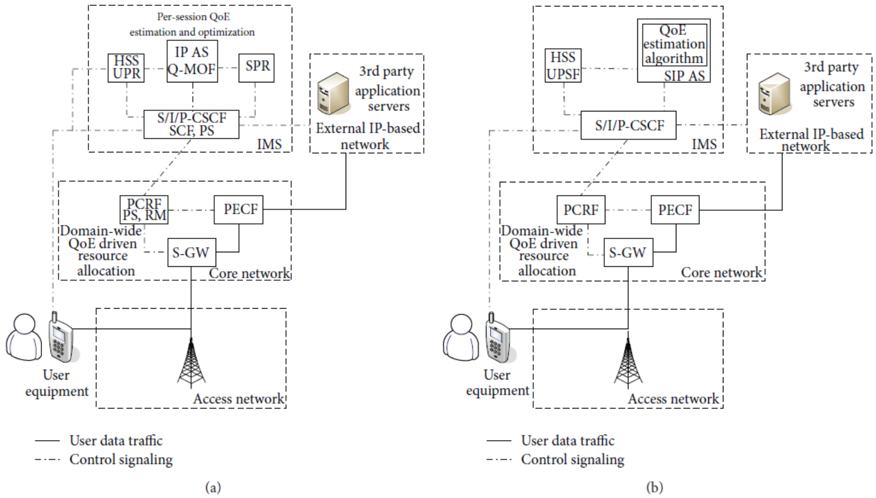


FIGURE 5: QoE control mechanisms: (a) Skorin-Kapov and Matijašević [100] and Ivešić et al. [101]; (b) Sterle et al. [102] (see Abbreviations).

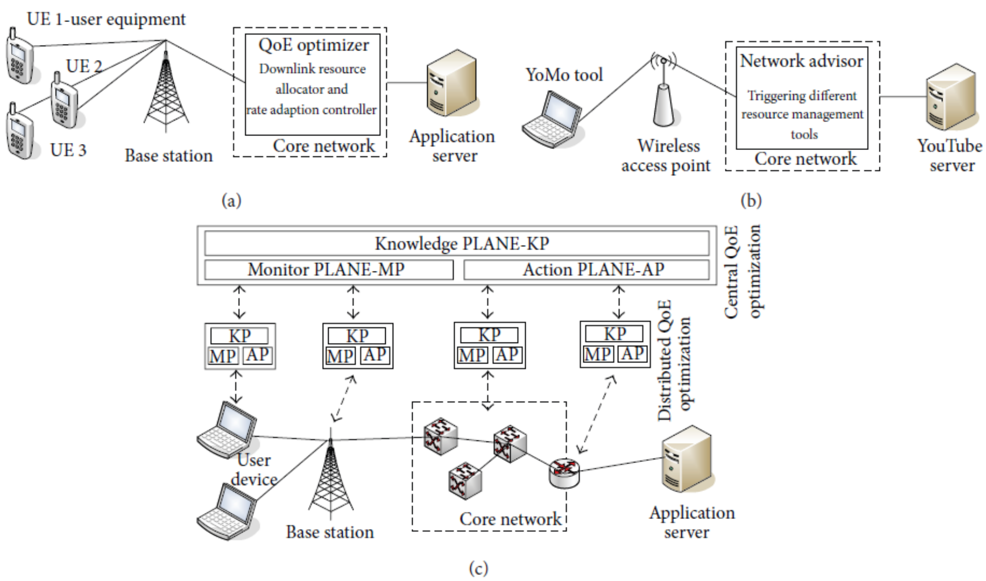


FIGURE 6: QoE optimization mechanisms: (a) Thakolsri et al. [103]; (b) Staehle et al. [97]; (c) Latré et al. [104].

4.6.2 User QoE as a Marketing & Customer Care Tool

4.6.2.1 Story Highlights

In today's networks, analysis of quality is often performed on a gross level – e.g., assessing coverage by looking at received signal quality at sampling points over a particular area. However, individual users are not able to perceive these gross indicators. The experience of each individual user is determined by how their applications perform on their devices under their usage patterns and at their locations. Collective measurements in system performance don't reflect the true QoE delivered to individual users and nor do they allow differentiation of users based on their usage patterns or commercial importance. 5G networks should allow the measurement and management of user QoE as a tool to improve customer care and provide marketing points for the network. The following use case illustrates how this capability may be used:

Michelle is a high value user who uses her mobile extensively for work and pleasure. Michelle's home is located in an area of generally good coverage but due to local effects at her house the indoor performance she achieves is patchy and inconsistent. This is reflected in stalls and glitches when Michelle attempts to do video conferencing or video streaming. The 5G system monitors the QoE Michelle receives and algorithmically determines that, at the application level, she has a poor QoE for video services. From the packet level this may be manifested by high amounts of jitter in packet delivery times. Alternatively, the application itself may collaborate in reporting the QoE to the 5G network. Based on the poor QoE, a customer care intervention is automatically triggered and Michelle is offered a home base station as part of her existing package to address her particular user experience at home.

Michelle works in a downtown office in an area where network load is very high during office hours. This means that when at work Michelle is often using a congested mobile network. This noticeably degrades her experience when using cloud based productivity applications on her mobile. System QoE monitoring tools detect that Michelle's interactive applications are performing poorly. Based on her profile as a high value customer and the interactive nature of her applications the network automatically reprioritizes Michelle's traffic so that her interactive application use receives higher priority than non-interactive downloads from other users and services to users with lower commercial value.

4.6.2.2 Business Drivers

Accurate measurement and management of QoE, as experienced by actual users, allows networks to adapt services to the personal requirements of users. When the network is congested priority can be given to those applications and users which have the highest commercial value and most overall benefit. Factors that cause a degraded service for users can be detected and addressed, even if these are not visible at the gross level.

By managing user QoE, networks can improve customer satisfaction, reduce user churn and improve their value to customers by tiering their services.

QoE measurements can also improve the customer service experience if customers contact the network operator to report problems by allowing the operator to quantifiably assess problems that the customer may be experiencing. This data will help translate the customer's subjective experience into actionable problem solutions.

4.6.3 Analytics to Predict QoE & Required Network Enhancements

4.6.3.1 Story Highlights

Today’s network capacity and performance engineering largely focuses on determining where network investments are needed to improve coverage, signaling capacity, and broadband throughput. By moving toward an approach that includes measurement of End-User QoE as a primary driver of investments in Network Enhancements, we can better direct such investments. It also becomes possible to predict End-User QoE performance based upon traffic growth and other trends.

Measuring and using End-User QoE as a Traffic Engineering parameter would involve:

- Understanding high-value users' most important applications including changes in these users' application mix.
- Identifying where in the network these users want access to important applications.
- Being able to properly assess high-value applications performance when throughput, latency, coverage, or other network performance measures are operating at different levels.
- To accurately determine the impacts in the previous bullet, one must look at network performance on an end-to-end basis including elements such as:
    - Backbone connection performance.
    - Edge and core cache effectiveness.
    - Performance of backhaul and other connectivity elements.
- Measurements to trend changes in high-value users' QoE based upon:
    - Likely growth in overall network load over time.
    - Location from which the network is accessed.
    - Changes in traffic pattern of important applications based upon application upgrades, etc.

4.6.3.2 Business Drivers

- Better return on network investment due to more direct application of network investments specifically targeted to improve QoE for high value users.
- Better retention of high-value users based upon direct measurement and trending of their QoE.
- Application centric end-to-end approach based upon QoE can unlock more effective network upgrade strategies. For example, understanding that a specific video app and the associated high value users might directly benefit from improved QoE could drive investments in edge caching at network locations that are frequented by the target users vs. more broadly applied upgrades to increase network bandwidth.
- Ability to project changes in end-user QoE enables a range of customer care and management options that are difficult to realize with the traffic trend data that is commonly used today.

4.6.3.3 Deployment Model

- Mechanisms are required to understand end-users' high-priority applications and measure their performance. This will require a mix of network and end-user device measurements.
- Analytics will be needed to identify high-value groups of end-users and their usage patterns in terms of applications and network access locations.
- Traffic visualization and engineering tools and processes will need to be updated to use information gained via the mechanisms in the previous two bullets to predict QoE trends and the impacts of potential investments in network enhancements on these trends.

4.6.4 Device-Centric Approaches to QoE Use Cases

4.6.4.1 Story Highlights

The breakthrough in this feature is if a user were to start to select a certain application, the network would be advised and may move the user to a more appropriate RF channel or cell. For example, say a user is currently on a macro cell and starts to select a streaming video service. Before that service is even launched, the UE could advise the network that it is going to need a certain QoE. The network could then handover the UE to a small cell site or add a carrier that can provide the needed data bandwidth to the user. In this way the service is launched on the best available part of the network. The network is not overburdened providing service in the least appropriate RF channel or cell site.

4.6.4.2 Business Drivers

Improved customer satisfaction.

4.6.4.3 Deployment Model

Instead of a meaningless signal strength meter the device can display what the network is capable of handling. For example, 'HDVideo-Streaming' as an icon when there is unrestricted capacity available. This would work in concert with network analytics.

When less capability is available on the network the device may only display an icon indicating 'VOICE' or 'TEXT' services.

If a user begins to select a service for which the network does not have adequate resources, prior to the resources being requested the UE could be told to possibly select Wi-Fi if it is available or that it will experience a degraded experience.

If the network cannot handle UL requests of large files such as photos, the device should indicate to the user that uploads are suspended until the network regains capacity. This will reduce user frustration when users think that a file has been uploaded but are unable to find it and resend it again, which uses up even scarcer network resources.

4.6.5 Augmented Reality

4.6.5.1 Story Highlights

Augmented reality (AR) is the process of augmenting (superimposing) graphics, images, audio, or other information over a live direct or indirect view of a physical, real-world environment.

There are a number of applications that may benefit from AR: entertainment, medical, education, commerce, public safety, and many others. An example of AR in a recent popular video game *Pokémon Go* is shown in Figure 4.1.

Figure 4.1: Pokemon Go Augmented Reality

Figure 4.2 shows another example of AR in the field of medicine.

Figure 4.2: Augmented Reality to Improve Visualization of Medical Situations

Components of AR generally include a display (hand-held, head-mounted, projection, contact lens, etc.), a tracking system (digital camera, optical sensor, gyroscopes, accelerometers, GPS, etc.), and the relevant processing hardware and software.

4.6.5.2 Business Drivers

AR provides an enhanced user experience. It has already made its way in many devices such as automotive, video games, and others. It is projected that AR will have $90B in revenue by 2020[9]. Facebook recently paid $2B for Oculus. Magic Leap took $1.4B from Google (and others). Apple recently acquired Metaio and declared AR as "a 'core technology' for the company"[10].

4.6.5.3 Deployment Model

1. A user walks through a history museum wearing a head-mounted AR device. As she stands in front of some of the exhibits, additional information related to the item is presented to her on the head-mounted display. Some of the item may even "come to life" in motion or provide an "X-Ray" view of the internal mechanism.
2. Public Safety dispatchers use head-mounted AR to get real-time information about their surroundings, including the location of other dispatchers and their vital signs, as well as potential areas that may need additional support or should be avoided.
3. A surgeon wearing a head-mounted display can "see" hidden parts of the body, and thus avoids hitting a major artery or nerve during a surgery.
4. A shopper holds her phone in front of store fronts in a mall to see details of what each store offers, any specials, coupons, etc.
5. Ikea provides an app that will allow a family to see how a sofa will look like in a house before it is purchased and delivered.

4.6.6 Virtual Reality

4.6.6.1 Story Highlights

Virtual reality (VR) is the process of generating realistic images, sounds, and other sensations that replicate a real environment. It simulates a user's physical presence in that environment by enabling the user to interact with the space and any objects depicted therein using specialized display screens or projectors and other devices.

There are a number of applications that may benefit from VR: gaming, entertainment, training, education, therapy, and many others.

The main component of VR is a headset, also called Head-Mounted Display (HMD), with sensors detecting user's motion and synchronizing the virtual environment to it. As shown in Figure 4.3, some advanced VR systems may also include other components such as haptic systems for tactile information, multimode devices, and vibration feedback. The objective is to create a lifelike experience to the users by allowing them to interact with the virtual environment.

---

[9] < http://www.digi-capital.com/news/2016/04/the-reality-of-120-billion-arvr-business-models >

[10] < http://www.extremetech.com/computing/233735-apples-tim-cook-declares-augmented-reality-a-core-technology-for-the-company >

Figure 4.3: VR Headset and multimode devices

4.6.6.2 Business Drivers

By some estimates, the market for VR will be worth $150B within four years[11]. Recently Facebook paid $2B to acquire Oculus which was founded in June 2012 and develops HDM for video games. Some of the game device manufacturers view VR as the future of gaming (e.g., Sony's PlayStation VR).

4.6.6.3 Deployment Model

1. A user "walks" through a 3D reconstruction of the Dudley Castle in England as it was in 1550. This was an actual application of VR dated back in 1994.
2. A couple interested in purchasing a house in another city do a virtual walk through of several houses in the market using VR.
3. A public safety command center supervisor uses VR to review an incident scene.
4. Surgeons at UCLA are using VR to practice highly technical and sensitive surgeries before they operate.
5. A patient uses VR as a therapeutic tool to receive various forms of exposure therapy, including phobia treatments.
6. Several friends play a group virtual game in a park with HMDs that are equipped with cellular handsets. The high bandwidth and low latency features of the 5G connection enhances the gaming experience.

[11] "The Race to Make Virtual Reality an Actual (Business) Reality" available at < http://fortune.com/virtual-reality-business >.

5 Network & Architecture Aspects

The definition of new radio access technology is a core attribute of 5G, but it is not the only important consideration. The development of the complete 5G system will impact the functionality and architecture of the core network as well. This section provides an initial consideration of the potential network and architecture aspects of 5G.

5.1 Network Slicing

5.1.1 Discussion

The use cases outlined in this white paper demonstrate that 5G networks will be expected to support services with very different requirements in a number of respects, including mobility support, availability, connection speed, reachability, power consumption, and complexity. For example, some devices will be accessing the network while moving at high speed, while other devices are expected to follow nomadic patterns or will be entirely stationary when accessing the network. In other cases, devices may only need a small portion of the functionality offered by the network; a device that automatically reports alarms, but does not support polling might not need to be reachable by other devices.

5.1.2 Network Slicing Description

A wide range of new diverse use cases will need to be supported by the 3GPP ecosystem. This needs to be done at the same time as continuing to support the traditional mobile broadband use cases. The new uses cases are expected to come with a wide variety of requirements on the network. For example, there will be different requirements on functionality such as charging, policy control, security, mobility, speed, availability, etc. Some use cases such as Mobile Broadband (MBB) may require application specific charging and policy control while other use cases can efficiently be handled with simpler charging or policies. The use cases will also have huge differences in performance requirements. Finally, the use cases will have requirements related to power consumption and complexity, meaning that it will not be acceptable to simply provide applications with the superset of all requirements.

In order to handle the multitude of segments and verticals in a robust way, there is also a need to isolate the different segments from each other. For example, a scenario where a huge number of electricity meters are misbehaving in the network should not negatively impact the MBB users or the health and safety applications. In addition, with new verticals supported by the 3GPP community, there will also be a need for independent management and orchestration of segments, as well as providing analytics and service exposure functionality that is tailored to each vertical's or segment's need. The functionality should not be restricted to providing isolation between different segments but also allow an operator to deploy multiple instances of the same network partition.

The figure below provides a high level illustration of the concept. A slice is composed of a collection of logical network functions that support the communication service requirements of particular use case(s). Some slices will be very rich in functionality, while other slices will be very minimalist, but the network slices are not arranged in any form of hierarchy. A given network slice will simply contain the functions required for a given application or class of service(s). Devices can be directed to the appropriate network slice in a way that fulfils operator and user needs, e.g., based on subscription or device type. The network slicing primarily targets a partition of the core network, but it is not excluded that the RAN may need specific functionality to support multiple slices or even partitioning of resources for different network slices.

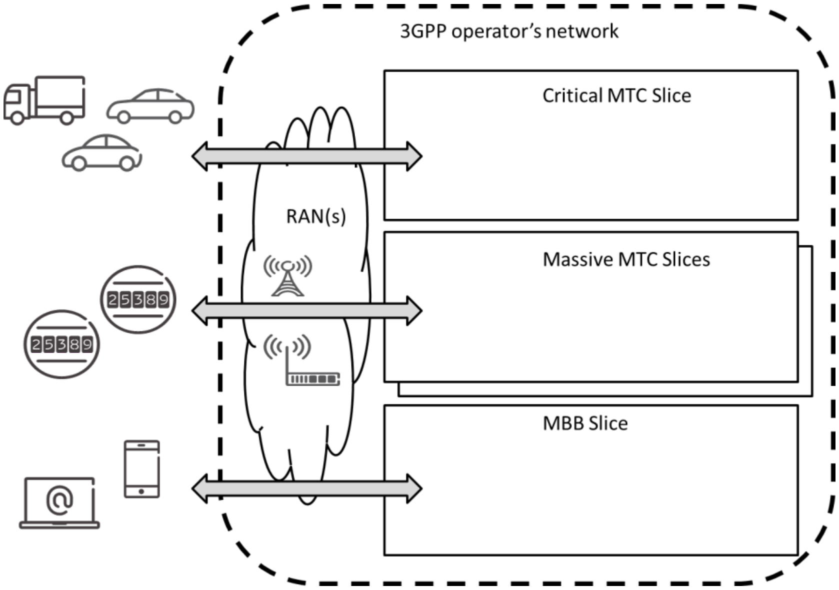


Figure 5.1: Network Slices that Cater for Different Use Cases

With the move towards network slicing for 5G services, the security requirements should take into consideration the nature of a multi-stakeholder environment proposed for Network Slices, isolation requirements between slices, and the specific individual needs of each stakeholder, whether it be the hosting platform provider or Network Slice tenant, in ensuring the integrity of the data processing and data held in the Network Slices.

5.2 Mobility on Demand

5G is expected to address a large variety of requirements for mobility support. Some mobile devices will be accessing the network while moving at very high speeds while other devices are expected to follow nomadic patterns or be entirely stationary when accessing the network.

5G is expected to see an explosion of IoT devices. These devices will require diverse mobility requirements driven by the IoT application. A variety of IoT applications leveraging cellular infrastructure can be envisioned; opportunities exist from power meters used in Smart Grid to Public Warning Systems utilizing wirelessly connected earthquake/tsunami detection sensors. All such applications can and are beginning to be deployed even on today's cellular networks. However IoT applications are predicted to grow at a much faster pace than perhaps what existing networks and cellular technologies can optimally handle. To support billions of IoT devices, a wireless network infrastructure is needed which is not only highly scalable in terms of its capacity, but can also optimally handle differing service needs of various IoT verticals. Examples of differing service needs include different requirements on mobility, as well as latency, network reliability, and resiliency. This diverse set of requirements may require re-architecting key components of the cellular network for example to support mobility-on-demand whereby mobility is only provided to those devices and service that need it.

The presence of a very large number of IoT devices on the cellular network could require radically new technology and solutions that the current 4G networks may not support. A key evolution needed to support mass scale IoT deployments is the scalability on the device as well as the infrastructure side. Current LTE networks are designed in a manner that makes such scalability technically and financially prohibitive. 5G systems must be designed to support connectivity to a massively larger number of low-cost, low-power, simple devices and device types in the "Internet of Things".

A key requirement of IoT communication is the low power consumption while still maintaining a high degree of reliability and coverage. Mobility-on-demand provides a selection of options, which may be dynamically assigned to a device or application according to the device and application context, or statically configured for specialized devices and applications. Mobility-on-demand consists of two components: one for managing mobility of active devices and a second for tracking and reaching devices that support a power-saving idle mode.

At the same time, requirements on mobility support also vary based on the applications and services used. While some services require the network to hide mobility events from the application layer to avoid interruptions in service delivery, other applications have application specific means to ensure service continuity. Hiding mobility events includes aspects such as minimizing interruption time and packet loss or maintaining the same IP address during intra- or inter-RAT cell changes.

It is worth noting that different applications may require different levels of mobility hiding support. While for instance some applications may not require the network to maintain the same IP address during mobility events, the applications may however still require the network to minimize interruption times so that they can continue to communicate quickly to ensure that their application-specific means of addressing mobility events work effectively.

It is also important to emphasize that application requirements on mobility hiding need to be considered in conjunction with the typical mobility patterns related to that application. If an application is typically only used when the user is mobile within limited regions, then it is sufficient to hide mobility events within those regions only. Other applications in turn may require the network to hide mobility events from those applications across the entire PLMN.

5.3 Network Aspects of User Experience Use Cases

Section 4.6 describes a number of use cases where the network is tuned to improve the user experience of end users. Implementing these scenarios requires four key solution elements in the 5G system:

- End user opt-in and control
- Measurements
- Analytics
- Targeted actions

The figure below illustrates these points. The arrows in the center of the diagram show the rich exchange and integration of data between these elements based on open communications.

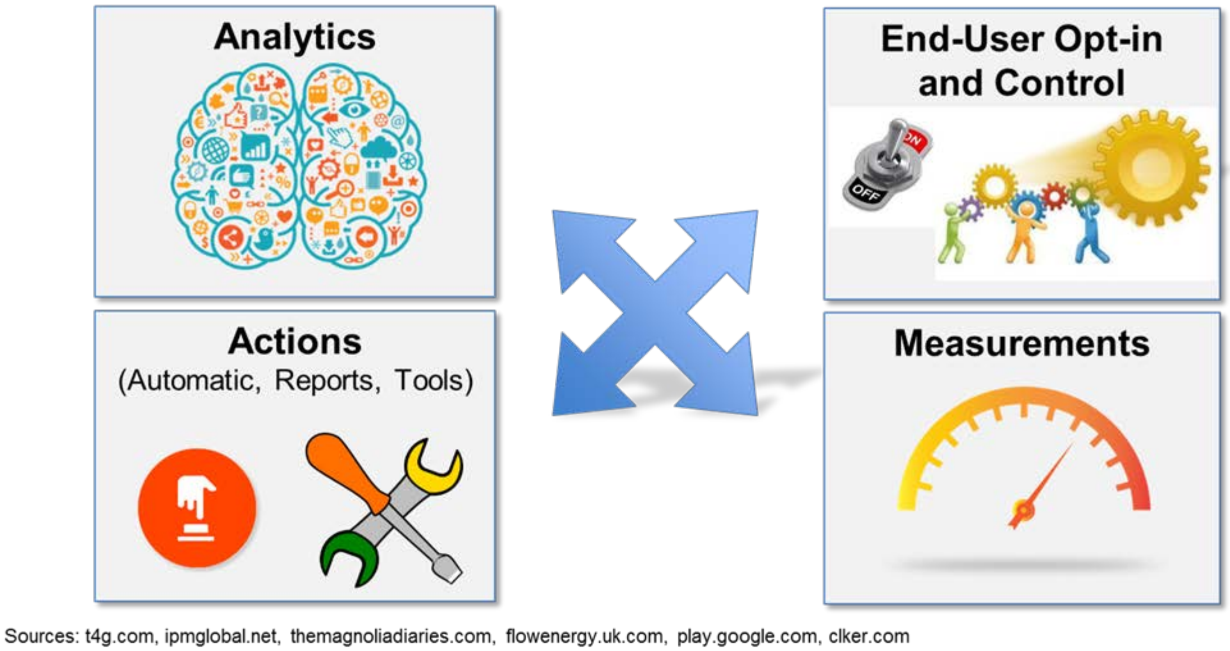


Figure 5.2: Solution Elements for User Experience Use Cases

A key element of improving user experience is to collect and analyze data for individual users in order to understand their service priorities and perception of the network service quality. In order to permit this data collection, the end user opt-in and control element will be provided for users to specify the aspects of service performance that are most important to them and to set their preferences for data collection in order to balance interests of privacy compared to the advantages of improved services. The ecosystem around user experience delivery will include third parties such as data brokers and content providers, and this must also be considered as part of the user control. Content providers will also need to opt in to collection of data on the usage of their applications and content.

The measurement element collects data from the network, content providers, and users in order to provide a factual basis for strategies to enhance user experience. Examples of data that may be collected include:

- Direct measurement of core network usage.
- RAN usage per cell.
- Historical usage data that can be combined with real time measurements of application usage and events to predict changes in traffic patterns.
- Feedback from users (crowdsourcing).
- Input from third-party applications (e.g., video streaming services).

The data collected through the measurement element must then be processed to extract actionable conclusions from the many data points. This is accommodated in the analytics element that applies data processing techniques to extract information about specific users or user groups from the available data. The analytics element should be based on what is termed a “workbench approach” to allow data analysts to define how they want measurement data to be processed, which factors are most important, and what information they wish to extract. The analysis may also be used to monitor the effects of actions taken and to assess their effectiveness.

The actions element is responsible for implementing or applying the actions determined by the analytics element across networks, backend and outboard systems (e.g., content caches, backend applications servers, etc.), and third party applications. This could be used to enhance or optimize third party services. It also provides capacity engineering tools and reports driven by the results of the analytics element for the purpose of optimizing network capacity investments to maximum overall delivered user experience.

6 Requirements

In Section 4 we looked at a number of use cases. In this section we consider potential requirements derived from those use cases; specifically, we consider the requirements that may be unique to North America. However, in order to identify such requirements, we must first study requirements being considered in other global organizations.

There are a number of global organizations that are studying potential requirements. As an example, Figure 6.1 shows enhancement of key capabilities from IMT-Advanced (“4G”) to IMT-2020 (“5G”)[12] being discussed in ITU[13].

[12] Figures reproduced with permission from the ITU. Source: Draft New Rec. ITU-R M.[IMT.VISION], ITU-R IMT Vision -“Framework and overall objectives of the future development of IMT for 2020 and beyond” developed by ITU-R Working Party 5D.

[13] The values in the Figure 6.1 are targets for research and investigation for IMT-2020 and may be further developed in other ITU-R Recommendations, and may be revised in the light of future studies.

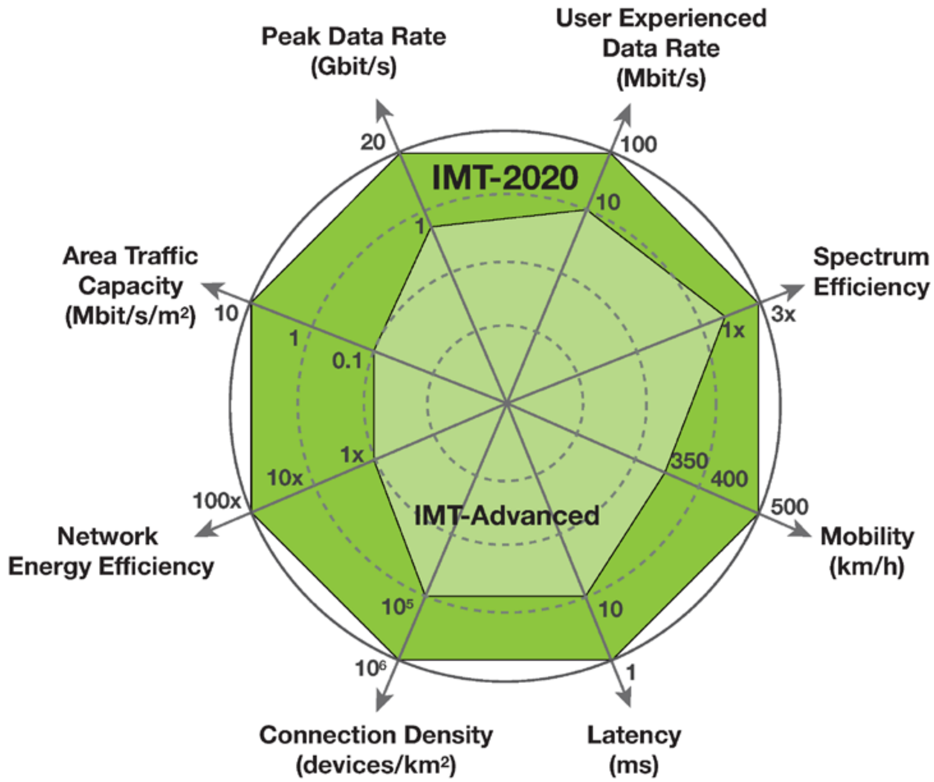


Figure 6.1: Enhancement of Key Capabilities from IMT-Advanced to IMT-2020

As the figure shows, there are several enhancements being considered to enable use cases in the next 5-10 years. For example, the peak data rate for enhanced Mobile Broadband is expected to reach 10 Gbit/s (20 Gbit/s peak data rate under certain conditions). This will enable several of the NA use cases, such as mHealth and Telematics that were previously discussed. Another example is lower latency that is important to the critical communication use case, or connectivity density that is crucial in IoT.

The majority of the requirements derived from use cases in NA are aligned with the global requirements. In general, the source of requirements that are unique to North America will fall into three broad categories:

- Regulatory requirements.
- Coherent network evolution.
- Security by design.

An example of the regulatory requirements is the performance criteria for public safety communications with Fire, EMS, Police and Emergency Operations Centers (FEPE) generally requiring high availability, low latency, and automatic-switched redundancy. In terms of network evolution, important considerations include the ability to remotely monitor end-to-end status in real-time, prevention of unauthorized access through data encryption or other secure solutions to protect emergency operations. It is recognized that the evolution of public safety communications has already begun and that sharing of best practices adopted by emergency management operations would be very beneficial toward advancing public safety needs.

The migration of public safety applications will not only be influenced by the current view of roadmaps, but also the future aspects of network evolution. In addition to the obvious technology innovations around new opportunities like Internet of Things, cloud applications, and software defined networking, 5G Evolution will be an important concern for PSRA applications. It will be important to identify unique North American PSRA marketplace requirements for Cellular-based Critical Communications. While 5G is still under development by the industry, there do appear to be significant opportunities in the future to address evolving public safety needs. These may include enhancements to redundancy and resiliency, security across networks, and capacity management during planned events and emergency response situations. As a result, the available roadmap of solutions for the public safety sector will evolve as the industry introduces new capabilities enabled by 5G developments.

In addition, Cybersecurity will be an important consideration as it intersects with network evolution and new user capabilities. Given the key needs of the public safety sector relative to secure and reliable communications and data transmission, it is anticipated there will be a need for effective solutions that can contribute to a longer term vision of how developments like network virtualization, software defined networking, cloud architectures, and enhanced end user control can contribute innovative cybersecurity solutions in the future.

The coherent network evolution requirement includes:

- Backward compatibility and interoperability below 6 GHz.
- Optimization above 6 GHz, which may compromise backward compatibility.
- Policy driven configuration and functionality.
- Mobility on Demand.

Finally, as mentioned earlier, security is an important requirement in NA. We believe security should be by design and not an after-thought.

6.1 AR and VR Network Performance Requirements

Augmented Reality (AR) and Virtual Reality (VR) are going to be important applications in the 5G timeframe. This document assesses the 5G radio and system performance requirements needed to deliver these AR and VR services.

6.1.1 AR and VR Implementation Options

The terms AR and VR refer to a particular type of user experience. In some cases, the boundary between these two use cases may be blurred and services may involve aspects of both AR and VR. These terms do not constrain implementations, and the AR and VR market may contain a wide variety of different implementation approaches.

One area where different implementations are possible is the location of technical functions and media content with respect to the cellular radio interfaces. Some implementations may make extensive use of content that is pre-loaded on the UE. Other implementations may rely more on real-time interactive content streaming from the network to the UE. Still other implementations may use a combination of pre-loaded UE content and real-time content streaming from the network. The type of design will depend on the service objectives and the operating constraints of the service.

The following sections discuss what level of performance may be needed to create an acceptable user experience for rich AR and VR services. Because different implementations are possible and services will vary in terms of their complexity, graphical sophistication and level of real-time user interaction there is not a simple answer to the question of the needed performance characteristics. Therefore, this section will outline some possible scenarios and discuss the implications.

6.1.2 Impact of User-Experienced Data Rate

The user-experienced data rate is a measure of the data rate that the 5G radio interface can deliver to a single user under reasonable operating conditions. This should be distinguished from the peak data rate which is only attainable under ideal conditions.

Some AR and VR services will have large amounts of content pre-loaded on to the UE and stored locally so that only the rendering of this content has to take place in real-time. In some cases, it may be possible to pre-load content over very high-speed WLANs or near-field radio interfaces. This type of interface can sustain a very high user data rate over a short range. In other cases, content may be slowly pre-loaded (e.g., during off-peak periods) to make use of spare radio interface capacity. These types of approach can help reduce congestion on the normal, wide-area, 5G radio.

For pre-loaded content, the user-experienced data rate is most relevant when the user wants to load content for later viewing and does not have access to a more specialized very high speed radio or the time to allow opportunistic pre-loading over a long period. In this case, the 5G radio should support a sufficiently high user-experienced data rate to allow the content to be pre-loaded in a reasonable time. Download sizes for high quality 3D games, which have many similarities with AR and

VR, are already multiple Gigabytes[14]; up to 100 Gigabytes is not unknown. Very rich VR and AR environments will tend to increase file size. To keep download times manageable (several minutes as opposed to hours), support for user experienced data rates of 100Mbit/s (750Mbyte/minute) and upwards are highly desirable.

Where content is not pre-loaded it may be streamed to the UE in real-time when it is needed (or just before). This case could occur if the content is so dynamic (e.g., because of the interaction of several users) or so extensive that pre-loading is not viable. The user-experienced data rate needed to support real-time content transfer will depend on the nature and quality of the virtual/augmented reality experience. Some augmented reality services may only add simple graphics on top of the real-world scene and may not need a very high data rate. However, real time delivery of very rich and realistic content will require a high user data rate (hundreds of Mb/s)[15].

6.1.3 Impact of Latency

When pre-loading content for later rendering, the latency does not need to be very tightly controlled as the data is predominantly transferred in one direction and not interactively. However, some content download may use congestion avoidance protocols that can have their performance degraded by high latency or high jitter. In this case the latency should be sufficiently low to allow the full user-experienced data rate to be utilized.

Where VR and AR content is being delivered in real-time or the environment is responding to real-time events, the latency can be extremely important. For example, in a multi-player shooting game all players need to have an accurate view of the environment, and when shots are fired they need to be immediately relayed to all players. In this case, and other cases, the latency needs to be considered in both the uplink and the downlink directions.

Another demanding case occurs if content needs to be streamed in real-time to adapt to changes in the viewpoint of the user (e.g., due to changes in user location in the real or virtual worlds or due to head movements). Low latency is needed to create a believable experience for these changes, particularly for head movements[16]. In order to support this, VR systems may use a higher frame rate than conventional monitors – e.g., 90 frames per second is used in some commercial VR systems. The requirement to very accurately synchronize movements and the displayed environment is sometimes called "the tactile Internet"[17]. Delays of less than 10ms between movements and the corresponding visual changes may be necessary ([2] and [3]).

[14] See < http://www.pcgamer.com/the-problem-with-growing-download-sizes/ > and ARRIS quoted by onlinereporter.com < http://www.onlinereporter.com/2016/06/17/arris-gives-us-hint-bandwidth-requirements-vr/ >.

[15] Cablelabs quoted by Variety < http://variety.com/2015/digital/news/looking-beyond-tv-cable-is-exploring-virtual-reality-1201571648/ >.

[16] See < http://blogs.valvesoftware.com/abrash/latency-the-sine-qua-non-of-ar-and-vr/ >.

[17] ITU Technology Watch Report "The Tactile Internet" < http://www.itu.int/oth/T2301000023/en >.

6.2 Mapping of 5G Capabilities to Use Case Requirements

The metrics and indicative values from Figure 6.1 can be used together with the use-cases described in section 4 to understand which use cases most benefit from the technical improvements and therefore may act as drivers for 5G roll-out.

The following table show how different performance metrics relate to important use cases. Note that the values shown are indicative of the ambition for IMT-2020 and not normative requirements.

Table 6.1: Performance Metrics and Associated Use Cases

Metric	**Indicative values**	**Use Cases**
User Experienced Data Rate (Mbit/s)	Now: 10 5G: 100	We understand the term "user experienced data rate" to refer to the rate that users may typically experience in a moderately loaded network under real-world operating conditions. This contrasts with the "peak data rate" which represents a very ideal performance situation. **Mobile Infotainment** Video streaming services up to HD quality can be supported in existing cellular systems subject to capacity limitations. For example, Netflix recommends a 3Mbit/s connection speed for SD video and a 5Mbit/s connection speed for HD video. The mobile Infotainment use case includes applications that have higher bitrate requirements such as 4k (Ultra-HD) video streaming. It has been estimated that 4k video streaming requires 10-12Mbit/s, and Netflix recommends a 25Mbit/s Internet connection. This performance may not be possible in real-world scenarios with current cellular technology. Moving from HD video to 4k video implies a significant increase in video bit rate requirements. Any significant user adoption of 4k video streaming and more demanding video applications will require 5G RATs.

Metric	Indicative values	Use Cases
User Experienced Data Rate (Mbit/s) (Continued)		**Virtual Reality and Augmented Reality** As discussed in section 6.1.2, Virtual Reality and Augmented Reality services may require a high user-experienced data rate. Support for services that require transfer of large amounts of content data also highlights the need to support efficient methods of content distribution throughout the system and to have content stored locally to the viewer as much as possible. **Public Safety Data/Video Service** Some public safety applications may require fairly limited amounts of data or low resolution video (e.g., for real-time viewing of body cams). Other applications (e.g., assessment of a disaster situation from a fixed camera) may benefit from high resolution video in order to allow viewers to assess details of a broader scene. Bit rate requirements may be similar to those addressed in mobile infotainment.
Spectrum Efficiency	Now: 1x 5G: 3x	Further spectrum for licensed and unlicensed public communication applications will likely be authorized in the 5G timeframe. However, not all licensed operators will necessarily increase their spectrum holdings. Not all new spectrum, particularly at high frequencies, will be suitable for all geographic situations. Spectrum for public communication services (both licensed and unlicensed) will continue to be a limited resource. Improving operational spectrum efficiency by increasing the density of base stations is expensive and not always possible. Improvements in spectrum efficiency due to the technical design of the RAT are necessary to economically enable increases in area traffic capacity. In terms of use cases, the major drivers to improve spectrum efficiency may arise as a result of two, related, causes: 1) Increased numbers of users using high bandwidth services. 2) Increased per-user consumption of bandwidth (e.g., the transition from HD to 4k video services). From a use case perspective, the most likely use cases to exhibit these characteristics are *Mobile Infotainment* and *Virtual reality/Augmented Reality*. As identified for “user experienced data rate”, these use cases include high-bandwidth services. These use cases also feature services targeted for mass consumer adoption. Together these effects combine to create a situation where spectrum efficiency will become an important consideration.

Metric	Indicative values	Use Cases
Mobility (Maximum Device Speed) (km/h)	Now: 350 5G: 500	The maximum device speed supported by existing standards is already sufficient for automotive applications. The targeted increase for 5G may be motivated by considerations of requirements for very high speed trains and, perhaps, air-to-ground applications. Neither train communications nor air-to-ground is explicitly addressed in this paper, though other sources have mentioned air-to-ground in the context of public safety use cases.
Latency (ms)	Now: 10 5G: 1*	*This value is shown in the figure above but its meaning is not defined, nor is this an agreed target. The achievable latency for 5G will be limited by the speed of light and real-world transmission architectures. The lowest latency figures may only be achievable for traffic that is locally routed between devices in relatively close physical proximity. Still it is generally accepted that 5G RATs will support lower latency than previous generations. **Automotive** Low latency could be a requirement of some automotive applications – e.g., for real-time coordination of vehicle behavior to prevent road traffic accidents. **Smart Cities** Low latency control can be a requirement for the real-time control and protection of infrastructure. For example, preventing cascading failures in high voltage transmission lines requires very low latency control of switching equipment. **Virtual Reality and Augmented Reality** As discussed in section 6.1.3, Virtual Reality and Augmented Reality services may require support of low latencies. **Critical Communications** Successfully responding to environmental warnings originating from sensors carried by first responders may depend on very low latency of message delivery.

Metric	Indicative values	Use Cases
Connection Density (devices/km^2)	Now: 10^5 5G: 10^6	Much of the growth in devices will be due to IoT applications. ATIS identifies several IoT use cases, any of which have the potential to require an increase in the device density beyond what is practical today: • Automotive • mHealth and Telemedicine • Smart Cities • IoT Device Focused Services NOTE: Supporting this connection density should also address the problem of mass-registration events, e.g., during network restoration.
Network Energy Efficiency	Now: 1x 5G: 100x	Energy efficiency is an important consideration on both operating expense (opex) and environmental grounds, but it does not strongly link to a particular use case.
Area Traffic Capacity (Mbit/s/m^2)	Now: 0.1 5G: 10	Area traffic capacity and spectrum efficiency are closely related in terms of their impact on use cases. The *Mobile Infotainment* and *Virtual Reality/Augmented Reality* use cases, discussed under “spectrum efficiency”, also have a dependency on area traffic capacity for similar reasons.
Peak Data Rate (Gbit/s)	Now: 1 5G: 20	The “User-Experienced Data Rate” is a better guide to the constraint for real-world use cases then the “Peak Data Rate”. Typically, the peak data rate is only available under idealized conditions. However, for demonstration purposes it may be of interest to exploit the peak data rate even if it is not representative of real-world experience.

Based on this analysis, one driver for the deployment of 5G RATs will be user adoption of demanding services of the type described in the "mobile infotainment" and “augmented/virtual reality” use cases. Another driver, particularly related to latency and the device density, could be the various IoT use cases. Several features of 5G RATs are also useful for critical communication applications, though this, on its own, may not be a sufficiently big market to drive 5G deployment.

7 Security Considerations

Security remains a top-of-mind item for network operators and their enterprise customers, as the variety and scope of threats continue to proliferate with the move toward 5G. Carrier networks provide critical infrastructure and services that governments and businesses depend on to operate every day; in addition public users now expect secure communications. The move to 5G, cloud services, and cloud-based data storage has only heightened awareness of the potential vulnerability of applications and data within the network. It should be noted that the security considerations identified in this section assume the use of a carrier 5G or evolved LTE RAN. The security implications of using previous generations of wireless technology, or unlicensed spectrum technologies such as Wi-Fi, are for further study.

Security provides the foundation of service assurance. Miscreants and the threats that they impose against the networks used to deliver critical services continue to get smarter, more agile, and more destructive. Carrier networks continue to converge, making it more important to properly segment threats and vulnerabilities by domain, but examine the aggregate threat landscape at the same time. Examples of this include the evolved packet core where traditional and mobile services share an infrastructure leveraging the Carrier Data Center and Cloud for operational efficiency and also for service delivery. Architectural innovation introduced by NFV and SDN introduces a new set of threats and mitigation strategies and also introduces a new set of visibility and control elements to handle the evolved threats.

In order to properly secure the carrier cloud, two fundamental elements are applied: visibility and control. Visibility refers to the ability to see and correlate information from the carrier cloud to baseline proper behavior and then to measure deviation from that norm. Simply said, “If you can’t measure it, you can’t manage it.” Visibility is defined as everything a network operator needs to do in order to:

1. Properly baseline the network to understand how it behaves under “normal” operations.
2. Gather telemetry and connect that telemetry to network functions and capabilities.
3. Use the telemetry gathered, feed it to a real-time analytics platform and deliver a security outcome flagging something as “anomalous.”
4. Finally, use that security outcome to take action, one of which may be to elastically deliver a mitigation control in the right context at the right time to properly remediate the “anomaly or threat” in question.

Sources of visibility come from traditional network measurements (e.g., NetFlow, BGP-FlowSpec, OpenFlow flow records, etc.), but the need to measure all aspects of a flow, from all elements of the carrier cloud to the application, has changed what data is collected and where we get it. Assessment can be carried out in all layers of a communications stack. An example of the new visibility includes the use of application level probes that are synthetically generated, collate information from all layers of the communication stack, and travel through the network to get a clear picture of how an application is behaving. Once all of the telemetry is gathered, a security controller will analyze it and determine, based on policy, suggested mitigation and controls to be applied.

Control refers to the actions taken to mitigate an attack. Some controls are taken proactively while others are applied after an attack takes place. Depending upon the assessment some actions may be automatically applied while others may be elevated to operational personnel for further analysis. There are two types of attacks. Day zero attacks are threats that we don’t previously have a fingerprint for. Typically, deviations in known good behavior of the carrier cloud and applications that request service and state from it are identified by the security controller, and some action is then taken to mitigate the attack or to get additional visibility; an action is sometimes taken to properly identify the miscreant. Day one attacks are threats that have a signature or fingerprint available and, quite often, a mitigation strategy exists in advance to handle the attack. Controls take the form of modifications to the carrier cloud to apply quality of service changes in per hop behavior to minimize the impact of an attack and also take the form of physical and virtual security assets applied as close to the source of the threat as possible in order to minimize collateral damage.

The information we have as carrier cloud operators is vast. Innovation in the way that we apply the information we have, in a closed loop iterative process, is a recent innovation in threat visibility and mitigation. The three elements of the closed loop iterative process are: Policy, Analytics, and the Carrier Cloud. Operators can now apply innovative methods to correlate geo-location information to behavioral analytics, compare those against policy in the context of a threat to the Carrier Cloud, and ascertain the nature of that threat and what to do about it with far greater clarity.

The above analysis assumes that the 5G equipment is under full control by the network operator. The concept of Network-as-a-Service, and related as-a-service models, requires further consideration. The assumption that the network operator owns the equipment and it is located within the network operator's premises using network operator-controlled communications links shifts to a multi-stakeholder environment not in direct control by the network operator. In this case, the approach to visibility needs to be holistic taking into account not only cybersecurity aspects, but also the physical security aspects of the infrastructure components.

In particular, the network functions (under an NFV architecture) are hosted on equipment (servers) under the control of third parties and often may be shared with other stakeholders. This introduces the need for the visibility of the security (cyber and physical) of the underlying infrastructure (servers, etc.) by the network operator. The traditional network operator vendor equipment, which includes scrutiny of the security of the equipment (managed through close SLAs with vendors), is being replaced by servers and routers (leased from third party internet compute services), the selection of which is not under direct control of the network operator. A further issue that will need to be addressed in such a multi-stakeholder environment is compliance with regulatory or legal requirements (e.g., LAES).

In consideration of this anticipated architectural model of hosting an operator's network infrastructure functionality (NFV) on a carrier cloud, alternative trust models providing security visibility into a network comprised of distributed and heterogeneous devices rather than the traditional "circle of trust" model operated by network operators will need to change. The FCC has raised this as a requirement, which was included in the recently published technical report *5G Security Requirements* (ATIS-1000077, January 2017) under Recommendation #8.

Visibility and Control properly applied to the advanced threats of today offer the Carrier Cloud a level of protection. We must continue to evolve, grow, and get smarter to keep our networks safe and resilient in the time of attack.

Encryption is becoming a pervasive reality in the networks we run today and 5G will be no exception to that rule. It is critical to have visibility at all layers of a communications stack as well as at different points in the network, allocated equally as pervasive, so as to ensure that proper visibility and control mechanisms can still be leveraged in network operations. This span is from the mobile endpoint all the way into the critical infrastructure that operates the mobile networks today. There are some early options in the industry today to enable such visibility to flows using metadata tags as a part of that flow. Standardization of such an approach is slowly happening in the form of protocols like NSH (Network Service Header of the IETF). In order for 5G to be a success, this proposal asserts that "network as a sensor" be considered to offer both the visibility and the mitigation control (mitigate a threat as close to the source in the right context as possible, so as to minimize collateral damage) to adapt to today's growing cybersecurity threat landscape.

8 Conclusions/Recommendations

8.1 Coherent Evolution

Cellular technology, particularly 4G LTE, has been incredibly technically and commercially successful. It has supported tremendous expansion in the number of users and the range of services enabled. Smartphones and similar devices using the 3G and 4G LTE networks offer a rich multimedia experience including web browsing, gaming, audio and video services, Smart Grid applications, and more. This success should be continued in 5G to enable a growing range of new and innovative services. In the evolution to 5G, the investment in existing technologies, particularly LTE, should be leveraged.

This white paper has shown viable deployment scenarios that include deployment of 5G, initially in urban areas, in a symbiotic relationship with earlier technologies. The coherent evolution of 4G LTE should be fully realized in the design of 5G.

While we emphasize the importance of 4G LTE evolution we also recognize that 2G GSM technology is now reaching the end of its useful life. The coherent evolution plan for 2G will provide a roadmap to reduce dependency on this technology and eventually eliminate 2G support. This will provide users with benefits of enhanced security and capacity and operators with benefits of network simplification and operating cost reductions.

8.2 Coverage & Capacity

Users' growing reliance on mobile connectivity continues to present technical and economic challenges for the provision of adequate coverage and capacity. Increasing utilization of virtual reality and augmented reality services are examples of changes in user behavior that can generate significantly greater performance demands on the network. We have shown various scenarios for how 5G technology can enhance coverage and capacity beyond that achieved by today's networks. Key aspects of this are:

- Intrinsic capabilities of 5G to tackle "difficult" environments particularly urban centers and the interior of high density buildings.
- Flexible use of available spectrum and access technologies including Wi-Fi, LTE-U, and other technology in unlicensed bands.
- Use of spectrum above 5GHz to provide high capacity, short range connections.
- Flexible models of network and coverage sharing including "neutral hosting" of in building services.

We recognize that sustaining excellent coverage and capacity is not a static event but a continuous process. It must be possible to efficiently reconfigure 5G systems to account for special events including planned events such as sports games and unplanned events such as natural disasters.

For smart devices, coverage requirements include international mobility ("roaming") and support for services in moving environments such as cars, planes, and mass transit. 5G should improve the user and operational experience in these areas.

8.3 Enabling IoT, Critical Communications, & Emerging Applications

The typical usage model for the development of previous generations of cellular technologies has been centered on the phone or smartphone. This type of device will continue to be important in the 5G timeframe but we see that cellular technology is now required to efficiently support a range of other emerging applications.

The term Internet of Things (IoT) is widely mentioned as an umbrella name for a range of applications that involve devices that are not phone like. In this paper we have identified important, specific applications of the IoT concept including:

- Healthcare,
- Smart Grid, and
- Transport and smart cars.

5G should support the full range of IoT applications including very constrained, low data devices and sophisticated devices with requirements for large amounts of real time data. IoT is one area where direct communication between devices, rather than communication via the network, may be desirable.

Augmented and virtual reality services are a rapidly growing market driven by advances in device capabilities, consumer excitement about new user experiences, and a range of practical applications. These services are expected to be demanding in their requirements for high bandwidth (particularly for virtual reality) and low latency (particularly for augmented reality). Meeting user demand for these services is expected to be an important driver for 5G.

Another emerging application area for cellular is in public safety and critical communication for first responders and other users. 5G technology should maintain and extend the critical communication technology being developed as part of LTE. This area will also have specialized requirements for IoT devices and wearable technology.

8.4 Implementing 5G for Efficiency & Flexibility

Commercial drivers mean that 5G systems must make improvements in all aspects of operational costs and efficiency. This requires the 5G system to be more flexible than previous generations. In this paper we have identified the following concepts to show implementation approaches which can improve the flexibility of 5G systems:

- Mobility on Demand,
- Security by design, enabled by the “network as a sensor”, and
- Network Slicing.