October 6, 2014

VIA ELECTRONIC FILING

Ms. Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th Street, SW
Washington, DC 20554

Re: ***Protecting and Promoting the Open Internet,*** **GN Docket No. 14-28;**
Framework for Broadband Internet Service, **GN Docket No. 10-127**

Dear Ms. Dortch:

On October 2, 2014, Christopher Guttman-McCabe, Scott Bergmann and Krista Witanowski of CTIA – The Wireless Association® ("CTIA"); together with Dr. Jeffrey Reed, Willis G. Worcester Professor in Electrical and Computer Engineering and Director, Wireless@Virginia Tech, at Virginia Tech University met with Roger Sherman, Michael Janson, Chris Helzer, Jennifer Salhus, Jim Schlichting, Joel Taubenblatt (Wireless Telecommunications Bureau), Walter Johnston (Office of Engineering and Technology) and Scott Jordan (Office of Strategic Planning and Policy Analysis) of the Federal Communications Commission ("Commission").

CTIA explained the record in this proceeding makes clear that wireless is different, and highlighted three key factors that warrant a mobile specific approach: (1)Technical, operational, economic, developmental and other factors continue to make mobile wireless services different and warrant a mobile-specific regulatory approach; (2) competition and differentiation in the U.S. mobile broadband market are delivering an open mobile Internet, providing world-leading mobile broadband networks, and creating an innovative mobile ecosystem that benefits U.S. consumers and businesses; and (3) the FCC's 2010 open Internet rules have worked for consumers and the U.S. economy, with growth, innovation and investment rising significantly since 2010.

CTIA explained to staff that the Commission's tentative conclusion to extend a mobile-specific Open Internet framework is grounded in three aspects of the mobile marketplace: mobile broadband faces unique technical and operational constraints; mobile broadband technologies are rapidly evolving; and the "generally greater amount of consumer choice" for mobile broadband services than for fixed.[1] Dr. Reed reviewed with staff the attached technical paper he co-authored, *Net Neutrality and Technical Challenges of Mobile Broadband Networks*, and the attached presentation to help detail the operational constraints in these ever-evolving mobile networks, the complexity of mobile network management,

[1] Protecting and Promoting the Open Internet, *Notice of Proposed Rulemaking*, GN Docket No. 14-28, FCC 14-61, ¶ 91 (rel. May, 15, 2014); *see also id.* ¶ 62.

why flexibility is needed, and how prescriptive regulation would undermine mobile broadband operators' ability to provide consumers with the level of service they have come to expect.

We explained the primary technical factors affecting mobile network management; how mobile broadband providers apply differential treatment to different traffic streams on a real-time, dynamic basis; the stark technological differences between wireless and wireline networks and network management; and the problems that would arise from imposing prescriptive Open Internet regulation on mobile providers. The technical factors they highlight include the following:

- ***Scarcity of radio resources.*** With the explosion in the amount of mobile data traffic, spectrum resources have not kept pace. Mobile broadband operators are thus constrained, necessitating aggressive and efficient management of limited radio resources.

- ***Radio resource sharing.*** As the number of users being served by the same base station fluctuates, the challenge of providing high-quality service to each of them also grows, requiring providers to make choices regarding how to manage network resources.

- ***Dynamic channel conditions.*** The allocation of radio resources constantly changes due to changing channel conditions and the interference environment, as often as every millisecond.

- ***Varying resource consumption.*** For a given channel condition, different services consume different amounts of resources. Thus, resource allocations change as users shift among different uses – often many times during a given session.

- ***Integration of devices and the network.*** Even when two devices experience identical channel conditions and allocation of radio resources, their design characteristics may dictate widely different throughput, further complicating network management.

- ***Ever-evolving network.*** Mobile broadband providers constantly manage user mobility across various technology generations and revisions across the network, offering differing levels of achievable network performance.

- ***Challenges of network capacity additions***. The intricacies of capacity growth (adding spectrum and wireless infrastructure deployment), along with ever-rising user traffic, make efficient utilization of the existing radio resources extremely critical to the user experience and network efficiency.

We also explained that mobile and fixed networks face vastly different technical challenges, as compared with fixed broadband offerings. Fixed networks have significantly higher capacity and predictability of resource requirements, whereas mobile networks are far more capacity constrained, with constantly changing user requirements and operating environments. Fixed networks involve channels that are relatively clean with signal regeneration, while mobile channels are impaired with interference, multipath and blockage, varying by location and from one millisecond to the next.

We noted that mobile broadband providers need more flexibility to manage their networks and to ensure that their customers have the service they have come to expect. That flexibility must include the ability to manage applications to avoid harm to the network and to maintain reliable and efficient service for the aggregate user experience. Similarly, mobile operators should be free from any anti-discrimination or commercial reasonableness requirement that would restrict their ability to innovate, optimize, and differentiate service to deliver a high quality product. In addition, expanded transparency requirements are infeasible in the context of dynamic, ever-changing mobile network operations. We explained that more prescriptive mobile rules would only stifle innovation and competition, to the detriment of the user-experience and system capacity, and would severely limit the ability of mobile wireless networks to meet the unique challenges faced by modern wireless networks.

Dr. Reed detailed the difficulties that would be created by the application of an overly broad or overly prescriptive set of rules on mobile broadband and the inability to remedy overbroad rules through the application of a reasonable network management exception. As his paper explains, "subjecting this type of network and network management to broad prophylactic rules with a vague 'exception' standard would provide no clarity to carriers, edge providers, or consumers as to how these networks will be managed. The exception would either simply subsume any rules (e.g., blocking or non-discrimination) or providers would be stripped of their ability to evolve and manage networks for the betterment of the entire subscriber base."

Pursuant to Section 1.1206 of the Commission's rules, 47 C.F.R. § 1.1206, this letter is being electronically filed via ECFS with your office. Please direct any questions to the undersigned.

Sincerely,

/s/ *Scott Bergmann*

Scott Bergmann
Vice President – Regulatory Affairs
CTIA – The Wireless Association®

cc: Roger Sherman
Michael Janson
Chris Helzer
Jennifer Salhus
Jim Schlichting
Joel Taubenblatt

Walter Johnston
Scott Jordan

Attachments

Net Neutrality and Technical Challenges of Mobile Broadband Networks

Dr. Jeffrey H. Reed and Dr. Nishith D. Tripathi

September 4, 2014

TABLE OF CONTENTS

Abstract 1

1. Overview 3

2. Mobile Wireless Networks: Evolution, Network Architecture, and Operations 6

   2.1 Evolution of Mobile Wireless Networks 6

   2.2 Network Architecture 8

   2.3 Typical Wireless Network Operations 11

3. Characteristics of Mobile Wireless Networks and Differences Between Wireless Networks and Wireline Networks 12

   3.1 Characteristics of Mobile Wireless Networks and Resulting Implications 13

   3.2 Differences Between Wireline Networks and Mobile Wireless Networks 17

4. Challenges of Implementing the FCC's Proposed 2014 Rules on Net Neutrality to Mobile Wireless Networks 20

   4.1 Transparency Rule and Mobile Wireless Networks 20

      4.1.1 Network Management Practices 20

      4.1.2 User Experience and Network Performance 22

      4.1.3 Congestion 24

      4.1.4 Peak Load Management 25

      4.1.5 Parameters of Best-Effort Service 25

   4.2 No-Blocking Rule and Mobile Wireless Networks 27

   4.3 Unreasonable Discrimination Rule and Anti-Discrimination/Commercial Reasonableness Rule, and Mobile Wireless Networks 28

5. Mobile Wireless Broadband Internet Access: An Integrated Information Service 31

6. Recommendations 36

Appendix I 38

Figures/Tables

Figure 1. Ever-Changing Mobile Wireless Networks 7
Figure 2. Simplified 3G and 4G Network Architectures 9
Figure 3. Operations of the Wireless Mobile Device and the Network 11
Figure 4. Characteristics of Mobile Wireless Networks 13
Table 1. Summary of Differences Between Wireless and Wireline Networks 19
Figure 5. Necessity of User Discrimination Due to Dynamic Radio Environment 29
Figure 6. Major Communication Steps for Web Browsing 34
Figure 7. Major Communication Steps for Video Conferencing 35
Figure 8. Integrated Information Service Offered by the Mobile Broadband Network 38

Abstract

As this paper describes in detail, the management of mobile broadband networks is a constantly evolving task. From millisecond to millisecond, handsets with differing capabilities, consumers with different usage patterns, applications that utilize different aspects and capabilities of both the handset and the network, and content consumption, including video, must be integrated with the network and managed adroitly to deliver a world-class broadband experience for the customer. Now imagine that millisecond to millisecond process happening while the consumer is in motion, while the handsets vary in capability (think flip-phone to smartphone), while the available network changes from 3G to 4G and from one available spectrum band to another, while traffic moves into and out of a cell sector, and while spectrum capacity is limited. **This entire process – the integration of all of these different variables – is unique to mobile broadband.** This paper is designed to illustrate and explain this extremely complex, very dynamic process in the context of the FCC's 2014 Notice of Proposed Rulemaking ("NPRM") on net neutrality.

The NPRM seeks comment on several proposed rules and associated mechanisms. In particular, the NPRM seeks comment on three rules that impact the management of mobile broadband networks. First the "transparency rule" requires mobile broadband providers to publically disclose accurate information regarding network management practices, performance, and commercial terms of their broadband Internet access service. Second, the "no-blocking rule," which was vacated by the D.C. Circuit, prohibited mobile broadband providers from blocking consumers from accessing lawful websites, as well as prohibited blocking applications that compete with the provider's voice or video telephony services, subject to "reasonable network management." The Commission is now exploring modifications to these rules. And third, though it tentatively concludes that such a rule should not be imposed on mobile broadband providers, noting its previous findings distinguishing mobile broadband in the context of net neutrality regulation, the Commission also seeks comment on whether it should apply to mobile broadband networks an "anti-discrimination/commercial reasonableness" rule, that would enforce a "commercially reasonable" standard of conduct for broadband provider practices.

This paper demonstrates that any extensions of, or additions to, the FCC's 2010 rules would be unwieldy and over-inclusive when applied to the complex and constantly-evolving management of mobile broadband networks. In fact, with the introduction of LTE, networks are managed and operated in a far more complicated and complex manner than the networks in place in 2010 when the Open Internet Order was adopted. As more of the LTE standard's advanced functionalities are incorporated into wireless networks, the complexity and prioritization in the networks will only grow, as will the benefits to consumers.

This paper addresses, based on the complexity and constantly-evolving management of mobile networks, why several of the proposals could be disruptive to a robust consumer broadband experience, and why some of the Commission's tentative conclusions should be maintained. For example, requiring mobile broadband providers to develop and/or report metrics regarding network management would be extremely difficult from a technical perspective and is unlikely to be useful due to the millisecond-to-millisecond adjustments that are inherent to a mobile broadband network. As described throughout, ever-increasing usage and scarcity of spectrum resources requires active management of the network to address capacity issues in a rapid fashion at the cell (or sector) level based on the demands placed on

the network. Similarly, while the competitive pressures on wireless carriers make imposition of a no-blocking rule unnecessary, broad application of a rule could have a significant negative technical impact on wireless broadband networks.

This paper explains how wireless applications can consume very large quantities of bandwidth, potentially causing problems for the end user or for others nearby. Third-party mobile apps and services can also interfere with and undermine network performance, and wireless network operators must be permitted the flexibility to manage their networks to prevent these negative effects. The NPRM also seeks comment on the feasibility of defining a minimum level of service that broadband networks must provide, proposing several possible standards that could be used. As discussed, such standards cannot be readily quantified for mobile wireless networks given the millisecond-to-millisecond adjustments in the network and would prevent wireless network operators from using techniques critical to ensuring a robust user experience. Also, as handset technology, base station technology, network technology and application technology rapidly change, it is unclear what metrics and standards would apply universally over time to fairly judge capabilities or performance.

The paper also demonstrates that the NPRM's tentative conclusion that an "anti-discrimination/commercial reasonableness" rule need not apply to wireless is the correct one. Differentiation among users and user services is required to provide a satisfactory quality of service to consumers. This is due to the dynamic nature of the radio environment and the need to operate good scheduling algorithm designs in a wireless network that maximize network performance while providing a good user-perceived experience. It is also due to product differentiation within a competitive marketplace in terms of what devices, features, and services might be offered as part of a carrier's service plan.

Finally, the paper explains that without today's real-time sophisticated scheduling algorithms that support network management that enables the service operator to cost-effectively provide services to many users simultaneously, overall user experience and network throughput will suffer. Treating all users alike at all times will degrade network performance by driving delivery to the lowest common denominator, and make the network less efficient. Adapting delivery to the predicted data delivery performance based on dynamic radio channel assessments promotes more efficient performance overall, across all users, even though at any single moment a network's site will distinguish between users based on channel quality.

The paper concludes that if adopted or expanded, several of the rules proposed in the NPRM would place constraints on mobile wireless networks that would stifle innovation and competition, negatively impact the user-experience and system capacity, and severely limit the ability of mobile wireless networks to meet the unique challenges faced by modern wireless networks. The result, in turn, would be harm to wireless users – the very outcome the Commission seeks to prevent.

From an engineering perspective, the concept that a network management exception to Open Internet rules is sufficient to allow wireless networks to evolve and operate is nonsensical. A modern wireless network must be managed aggressively. It is not an exception, it is a daily reality. Subjecting this type of network and network management to broad prophylactic rules with a vague "exception" standard would provide no clarity to carriers, edge providers, or consumers as to how these networks will be managed. The exception would either simply subsume any rules (e.g., blocking or non-discrimination) or

providers would be stripped of their ability to evolve and manage networks for the betterment of the entire subscriber base.

This paper demonstrates the following:

- *Minimal regulatory constraints for mobile broadband networks would facilitate achieving higher spectral efficiency and improved user experience.*
- *Network Management is practiced extensively in mobile broadband networks and is critical for wireless operations.*
- *Preserving the ability for wireless carriers to block websites or applications as necessary for reasonable network management is important to avoid harm to the network or degradation and is critical to maintaining reliable and efficient service.*
- *Application of an anti-discrimination/commercial reasonableness rule to mobile broadband providers would hamper their ability to innovate, optimize, differentiate, and deliver high quality products and services.*
- *Expanding the transparency rule would increase costs and negatively impact network management option, but will not provide any meaningful benefit to consumers.*
- *Mobile broadband Internet Access service is an integrated information service due to the tight coupling between the device and the many network elements, needed for customized processing of different types of information, and the distributed nature of the complex wireless network.*

1. Overview

This technical paper demonstrates the unique technical aspects of wireless broadband networks that make the imposition of prescriptive net neutrality regulations highly problematic. Mobile broadband networks are highly dynamic, with constant changes in network standards, technology, and capacity needs. Mobile broadband operators are also managing their networks with limited spectrum resources, which must be managed actively and quickly to provide a high quality of service to consumers. As a result, wireless network management practices are necessarily complex. Further, congestion-related metrics are highly variable both temporally and spatially, and also change by the millisecond, making meaningful reporting impractical.

The 2014 Net Neutrality NPRM. With respect to mobile broadband service, the NPRM discusses the transparency rule, the no-blocking rule, and a revised anti-discrimination/commercial reasonableness rule. The existing transparency rule requires the service provider to disclose items such as network management practices and performance, though the FCC now seeks comment on whether and how to expand the transparency requirements for mobile wireless providers. The proposed no-blocking rule would prohibit mobile broadband service providers from blocking consumer's access to lawful websites and from blocking consumer's voice or video telephony applications that compete with mobile broadband service provider's services, though the NPRM seeks comment on whether to apply this rule more broadly to mobile wireless services. The NPRM proposes an anti-discrimination/commercial reasonableness rule that prohibits commercially unreasonable practices based on the totality of circumstances. The NPRM tentatively concludes that this rule should not be applied to mobile broadband service, but it seeks comment on whether to reverse that finding. Comments filed in response to the NPRM affirm the technical findings explained in this paper.

Mobile Wireless Networks Undergo Constant Technical Evolutions. Mobile wireless networks have evolved from first-generation analog systems to fourth-generation high-performance digital systems with multiple revisions within a given generation. These generations and revisions have widely different capabilities for both the networks and the mobile devices. Commercial mobile providers typically have multiple generations and revisions of generations simultaneously operating to serve legacy and new devices. Each time a new revision is introduced network management practices must change. The mobile broadband network and the mobile device perform numerous operations and interact with each other so that the end users have anytime and anywhere seamless communications experience. And the wireless industry has not reached the end of the road on innovation – the industry is already turning to the development of 5G technologies, injecting further complexity in the design and management of mobile wireless networks.

Mobile Wireless Networks Have Unique Technical Characteristics. The difficulty of quantifying guaranteed network performance and user experience is increased further due to the unique characteristics of mobile wireless networks. Examples of such characteristics include:

- scarcity of spectrum,
- dynamic radio channel conditions,
- the need to share radio resources among numerous users and user services with different Quality of Service (QoS) requirements,
- mobility,
- vast variability in loading due to both variations in user density per area and variations in usage and data rates,
- inherently complex process of network capacity growth, and
- integration of devices and network technologies with widely different data use and application capabilities.

These characteristics pose significant challenges to mobile wireless networks and make the imposition of the prescriptive net neutrality rules infeasible. In particular, determination of any reliable universal thresholds or metrics to quantify user experience or network performance is infeasible. Further, imposing such specific metrics would then distort optimization and would impose conditions that would degrade consumers' mobile experiences. Furthermore, mobile broadband providers need a high degree of flexibility to efficiently and effectively manage precious radio resources to ensure the best possible aggregate service experience for all subscribers.

QoS and the ability to treat different types of traffic differently based on their service needs are essential in a mobile network. In a mobile network, where the connectivity performance is not as stable as with a wired network, some services will simply not work well if they are not subjected to differentiated treatment. VoLTE is one example – it is meant to replace the traditional, circuit-switched phone service available on cellphones. Without prioritization of this traffic, the quality and reliability of the phone service would be severely impacted. Other future services such as LTE multicast have similar requirements. As new services are layered onto the networks, and historical separation of data and voice services vanishes the need to address QoS issues will only increase.

Wireless Operators Engage in Numerous Network Management Techniques. The network management practices in mobile wireless networks are extremely complex and consist of numerous

mechanisms that are distributed among various components (or nodes) throughout the wireless and core network. Examples of network management mechanisms include the scheduling algorithm for downlink and uplink resource allocation, the handover algorithm, the load balancing algorithm, handling of the connected mode-idle mode transitions, adaptation to the changing channel conditions, power control, and interference coordination. These network management mechanisms are proprietary and are key competitive differentiators. Providers continually refine their network management practices to dynamically reflect changes in network equipment, application demands, and consumer usage patterns. Indeed, the rapid evaluation of these practices may well mean that by the time a given practice is challenged and adjudicated the practice may no longer be in use. Hence, a mandate to fully disclose these mechanisms, or to impose sweeping no-blocking or anti-discrimination rules, would discourage innovations, violate intellectual property rights, and harm consumers.

Wireless Network Operators Make the Most of Scarce Spectrum Resources. Wireless providers need maximum flexibility in the management of their networks to make the best use of the *scarce radio spectrum* in the presence of exponentially rising data traffic. Due to the scarcity of spectrum, innovative, high-performance, and ever-evolving network management mechanisms are absolutely essential to the overall network performance and user experience. For example, wireless providers must take steps to contain data-intensive applications from flooding the network with excessive amounts of traffic that would degrade service for many users. Wireless network operators require the flexibility to fairly balance network performance and user performance among users, devices, user services, and overall services on the network.

Net Neutrality Regulation Imposes Numerous Unique Challenges on Wireless Networks. As this paper demonstrates, application of the 2014 NPRM's proposed enhanced transparency rule to mobile wireless networks is nearly impossible, as network management practices are highly complex and are constantly changing. Furthermore, flexibility with respect to network management is essential to enable continued innovation in this area and these characteristics counsel strongly against far-reaching no blocking or anti-discrimination rules. Indeed, application of the no-blocking rule, meanwhile, is infeasible as the Commission has defined a "minimum level of service" that is not possible to guarantee for mobile wireless networks. The revised anti-discrimination rule is not intended to be applicable to mobile broadband service, and the findings of this technical paper strongly support this FCC conclusion. The FCC should continue to distinguish between mobile and fixed broadband with respect to the "no discrimination rule" and "anti-discrimination/commercial reasonableness rule." The dynamic and resource constrained (and at times, congested) nature of mobile wireless networks requires differentiation among users and user services to ensure a high quality of network performance and a satisfactory user experience.

Mobile Broadband Internet Access is an Integrated Information Service. Mobile broadband service is a highly integrated service that enables a subscriber to access a variety of services at once. The Commission itself observed that wireless broadband Internet access service offers a single, integrated service to end users that inextricably combines the transmission of data with computer processing, information provision, and computer interactivity. This level of integration requires cross-layer optimization in the network to ensure optimal network performance. Without the flexibility to actively manage their networks, mobile broadband providers will not be able to deliver integrated services at the level of quality that consumers have come to expect.

Recommendations. Due to the challenges faced by mobile network operators, which are outlined below, this paper recommends that the Commission:

- recognize that mobile wireless networks must be treated differently from other communications networks,
- strive for minimal regulation of mobile wireless networks to promote continued innovation, and refrain from applying far-reaching no blocking rules or an anti-discrimination/commercially reasonable rule,
- grant to network providers maximum flexibility regarding the design, management, and optimization of networks to serve consumers,
- refrain from establishing minimum performance standards (or metrics) for wireless networks, as these are impractical to define or enforce in the face of spectrum scarcity and variability, and
- ensure that proprietary and competitive network optimization and management processes are respected, which will ensure continued innovation and differentiation.

Flexibility in tuning and adapting the network management mechanisms to the fast-paced technology evolution, implementation of new features and uncertainty regarding the requirements of emerging applications or services urge that the network management mechanisms in mobile wireless networks should not be subject to broad disclosure, sweeping no-blocking, or anti-discrimination requirements. In other words, these network management mechanisms are intended, by their very nature, to optimize the aggregate performance for the benefit of all users. A focus on specific metrics may work to the detriment of the aggregate network performance and user experience. Conversely, reporting aggregate metrics will not reveal meaningful insights into specific instances.

2. Mobile Wireless Networks: Evolution, Network Architecture, and Operations

In order to fully appreciate the complexities associated with managing a wireless network and the difficulty of imposing an inflexible net neutrality framework, it is helpful to have an understanding of the rapid evolution of wireless networks and technology as well as the underlying architecture. In the more than 30 years that the wireless service has been provided to consumers, there has been a near-constant evolution of the underlying network. Section 2.1 summarizes this evolution of commercial mobile wireless networks. Section 2.2 illustrates the network architecture for the most popular 4G standard – Long Term Evolution (LTE). The wireless network and the mobile station (referred to as the user equipment or UE, mobile device, or handset device) perform numerous operations and interact with each other so that end users have anytime and anywhere seamless communications experience, processes which are quite different from wireline systems. Section 2.3 provides a glimpse of such operations and interactions.

2.1 Evolution of Mobile Wireless Networks

Mobile wireless networks have evolved from the first generation (1G) to the fourth-generation (4G) in just about three decades. Numerous 1G systems were used throughout the globe. Advanced Mobile Phone System (AMPS) is an example of the 1G system in the U.S. First generation systems were analog (radio air interface) in nature and offered primarily voice services. First generation systems evolved to second-generation (2G) digital systems. The 2G systems provided better voice quality and higher capacity compared to the 1G systems. Global System for Mobile communications (GSM), Interim

Standard-54 (TDMA), and later Interim Standard- 95 (IS-95 or CDMAOne) are examples of 2G digital systems engineered primarily for voice services used in the U.S. These digital systems evolved to '2.5 G' systems to better support low data rate uses, including GPRS for GSM, IS-136/EDGE for TDMA, and CDMA 2000 1X for CDMA. Due to expanding needs for wireless data at higher rates, third generation standards for mobile wireless networks focused on supporting data services more efficiently separated from voice channels. The 3G systems include a packet-switched core network to facilitate Internet access. Universal Mobile Telecommunication System (UMTS), High Speed Packet Access (HSPA), and 1xEvolution-Data Optimized (1xEV-DO as a CDMA derivative) are examples of true 3G cellular systems. The 3G systems support peak user data rates on the order of few megabits per second (Mbps). Finally, fourth generation systems such as Long Term Evolution (LTE) were developed to provide higher data rates (e.g., many megabits per seconds) and higher spectral efficiency. In addition, LTE would allow both data and voice to be provided in an integrated fashion using Internet Protocol (IP) for transport, also known as VoIP (Voice over IP). LTE is currently being deployed in the U.S. and around the globe and is expected to be the most dominant wireless standard for the near term. Mobile wireless networks will continue to evolve—indeed providers are already working on 5G—with future generations of technologies bringing new capabilities and challenges. It is key that this evolution and innovation be able to progress unfettered by restrictive regulation. Figure 1 depicts the evolution of mobile wireless networks.

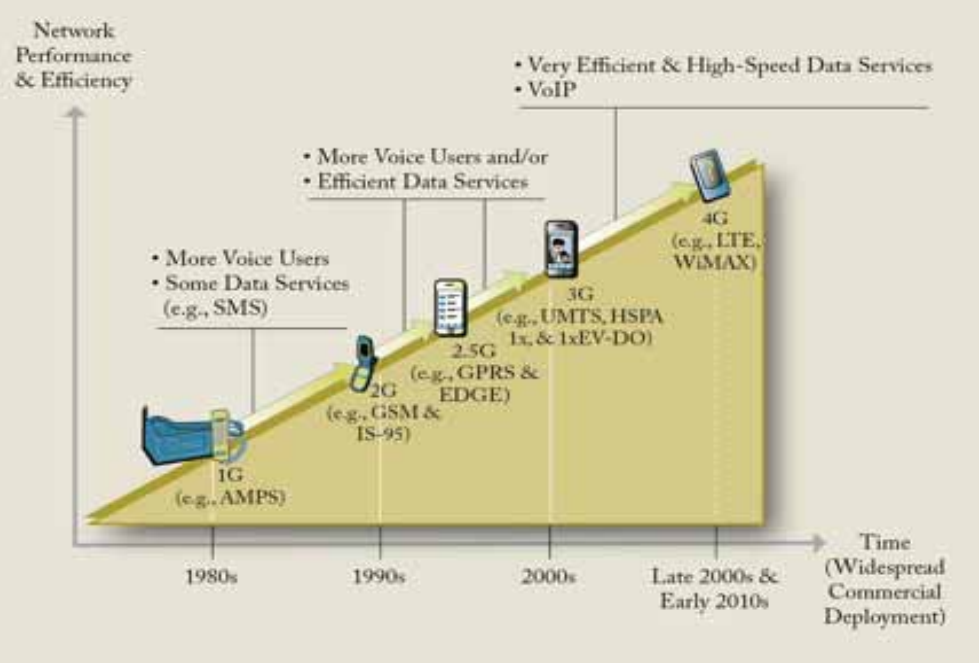

Figure 1. Ever-Changing Mobile Wireless Networks

Even for a given generation of wireless standards, multiple in-generation revisions that offer different features and capabilities exist. For example, 3G UMTS Release 99 supports a peak data rate of 2 Mbps in the downlink, while the 3G UMTS Release 5 feature called High-Speed Downlink Packet Access (HSDPA) supports a peak data rate of 14 Mbps in the downlink. The UMTS Release 7 feature called HSPA+ supports 21 or 42 Mbps in the downlink.

A mobile broadband provider typically has multiple revisions of multiple generations of technologies simultaneously operating. For example, in a given wireless service provider's network, some mobile devices may support GSM, some may support revisions up to HSDPA, some may support revisions up to HSPA+, and some may support revisions up to LTE. As the user switches from one generation of technology to another or from one revision to another, the performance can vary quite significantly. User mobility across different technologies needs to be properly managed by the mobile service provider. This involves complex network management.

The mobile service provider's network is never static. The network needs to be upgraded from one revision to another revision of a given generation technology and from one generation to another generation. Furthermore, once the network is upgraded with new features and capabilities, troubleshooting and then on-going optimization are carried out. The achievable peak performance keeps changing as the network undergoes never-ending upgrades. Even though LTE provides superior performance compared to prior generations of mobile wireless networks, LTE networks are currently undergoing upgrades with new features such as carrier aggregation and Voice over LTE (VoLTE), with each upgrade requiring changes to network management.

2.2 Network Architecture

The network architecture is different for 2G, 3G, and 4G (e.g., LTE) systems. This paper focuses on the network architecture for LTE due to its current dominance; however we will briefly describe simplified 3G and 4G network architectures below. In this section, we will describe the complex and decentralized nature of the wireless network and why application of net neutrality principles in this environment is so difficult. Moreover, with the move to an all IP-based infrastructure, the core wireless infrastructure is more intrinsically integrated into the radio network which in turn requires the wireless provider to calibrate and manage the radio resources and the core resources more carefully to ensure that subscribers are receiving an appropriate level of service.

LTE is defined by an organization or a standards body called the Third Generation Partnership Project (3GPP). 3GPP has defined a radio network called the Evolved-Universal Terrestrial Radio Access Network (E-UTRAN) and a core network called the Evolved Packet Core (EPC). The combination of the E-UTRAN and the EPC is termed Evolved Packet System (EPS) that can be viewed as the end-to-end LTE network. The LTE EPS uses the help of auxiliary networks such as IP Multimedia Subsystem (IMS) and the Policy and Charging Control (PCC) to provide a variety of services to end users. We will look at the main functions of the E-UTRAN, EPC, IMS, and PCC after a brief discussion of the simplified network architectures of 3G (e.g., UMTS) and 4G (e.g., LTE) network architectures illustrated in Figure 2.

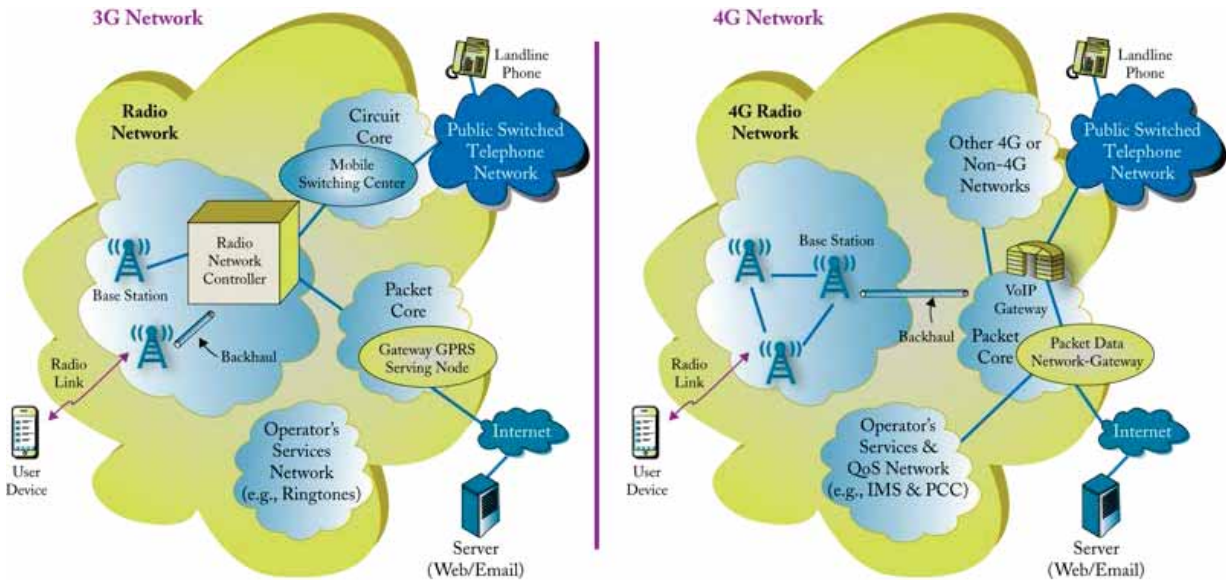

Figure 2. Simplified 3G and 4G Network Architectures

A 3G network consists of a radio network, a circuit-switched core network, a packet switched core network, and a services network. The radio network includes multiple Radio Network Controllers with each Radio Network Controller controlling hundreds of Base Stations. The Base Station communicates with the mobile device (referred to as the user device) via the air interface. The circuit-switched core network interfaces with the Public Switched Telephone Network so that the mobile device can communicate with a landline phone. The packet-switched core network enables the mobile device to access web and email servers via the Internet. The Mobile Switching Center is one of the nodes residing in the circuit-switched core network and controls the voice calls. The Gateway GPRS (General Packet Radio Service) Serving Node is an example of the packet-switched core network node and is in charge of assigning an IP address to the mobile device.

A generic 4G network consists of a radio network, a packet-switched core network, and a services and Quality of Service (QoS) network. The radio network includes the base stations. The packet switched network interfaces with the Internet using the help of a node such as the Packet Data Network Gateway. The packet-switched core network also interfaces with other 4G or non-4G networks. Since there is no circuit-switched core network in a typical 4G network, special nodes such as a VoIP gateway are needed to support calls between the 4G mobile device and the Public Switched Telephone Network. Auxiliary networks such as IP Multimedia Subsystem (IMS) and Policy and Charging Control (PCC) can be viewed as part of the operator's services and QoS network; these networks enable the service provider to offer to its subscribers a variety of IP-based services that have different QoS requirements. *We take a closer look at the LTE-specific 4G network architecture next.*

Wireless Radio Networks are Complex and Decentralized. The E-UTRAN has a decentralized and flat architecture. The E-UTRAN consists of the Evolved Node B (eNodeB or base station). The eNodeB communicates with mobiles over the wireless interface. The eNodeB makes the network management decisions related to the radio resource utilization. For example, the eNodeB evaluates the availability of the radio resources to determine if the subscriber can be offered services or not. The eNodeB implements a scheduling algorithm that allocates radio resources (radio bands and within one band,

Resource Blocks (RBs)) to the active users based on numerous factors including the target quality of service (QoS) of the applications of users, the amount of data, the number of users and the types of the user applications vying for resources, the radio channel conditions of users, the capabilities of the eNodeB and the mobiles, and the available spectrum. The eNodeB executes the scheduling algorithm as often as every 1 millisecond (ms). The eNodeB also determines the type of multiple antenna technique and the combination of the modulation and coding scheme for a given mobile device to reflect the prevailing radio channel conditions for the mobile. The eNodeB also carries out load balancing and interference coordination with the neighboring eNodeBs. The eNodeB implements a handover algorithm that utilizes the measurement reports of the radio environment received from the user equipments (UEs) and makes a handover decision if appropriate.

The Core Network is Tightly Integrated with the Radio Network. The Evolved Packet Core includes several entities such as the Mobility Management Entity (MME), Serving Gateway (S-GW), Packet Data Network Gateway (P-GW), and Home Subscriber Server (HSS) with specific responsibilities assigned to these entities. The Mobile Management Entity authenticates the LTE subscriber by working with the Home Subscriber Server. The Home Subscriber Server stores the subscriber database including the authentication related information. The Mobile Management Entity keeps track of the mobile device location when the mobile is in the idle mode so that a page can be sent to the mobile device to bring it out of the idle mode. The Mobile Management Entity coordinates the setup of Evolved Packet System bearers[1] for a mobile device; the Evolved Packet System bearers help carry the user traffic between the mobile and the Packet Data Network Gateway. The Packet Data Network Gateway allocates one or more IP addresses to the mobile device. The Packet Data Network Gateway is a mobile's gateway to the outside world such as the Internet. The Serving Gateway helps move the traffic between the eNodeB and the Packet Data Network Gateway. When the mobile goes from one eNodeB area to another eNodeB area, the Serving Gateway learns about such user mobility from the Mobile Management Entity and is able to forward the traffic between the Packet Data Network Gateway and the correct eNodeB. When the user is receiving information from a web server, the IP packets from the web server pass through the routers in the Internet and arrive at the Packet Data Network Gateway. The Packet Data Network Gateway forwards the user traffic to the correct Serving Gateway. The Serving Gateway forwards the IP traffic to the eNodeB that is currently serving the UE. The eNodeB allocates suitable radio resources to the mobile device and sends the IP packets to the mobile over the air interface.

The End-to-End LTE Network is Carefully Calibrated to Provide Quality of Service to Consumers. The Evolved Packet System works with the IP Multimedia Subsystem (IMS) and the Policy and Charging Control so that subscribers can be offered a variety of IP Multimedia Subsystem-based services with suitable QoS. The QoS benchmarks are derived from the standards work in 3GPP and are not set by the individual wireless provider. Examples of IMS-based services include Voice over IP (VoIP), Short Message Service (SMS), and Instant Messaging (IM). The wireless service provider is aware of the IMS-based services of the subscriber and the signaling associated with the IMS-based services passes through the Evolved Packet System and the IMS network. The IMS network performs its own service authentication for the cellular subscribers to allow the subscribed IMS services. The IMS network processes the signaling messages and extracts QoS for a given IMS service. The IMS network specifies

[1] A bearer in this context refers to a "pipeline" connecting two or more points in the communication system in which data traffic flows. An "EPS Bearer" would be the pipeline through which data traffic flows within the Evolved Packet System.

such QoS to the Policy and Charging Control network, which compares the service-requested QoS with the subscribed QoS and determines the QoS and charging rules based on operator policies and user subscriptions. The Policy and Charging Control network uses the help of the Packet Data Network Gateway to initiate the setup of an Evolved Packet System bearer[2] to meet the QoS requirements of the subscribed IMS service. Non-IMS services such as regular email and web browsing use the best-effort Evolved Packet System bearer toward the Internet, and signaling and traffic for such non-IMS services do not pass through the IMS network. Once a suitable Evolved Packet System bearer is in place, the Policy and Charging Control and the Packet Data Network Gateway implement the negotiated service-specific QoS. Although the resource bottleneck is usually radio resources at the eNodeB, the QoS control is needed on the link between the eNodeB and the Serving Gateway and the link between the Serving Gateway and the Packet Data Network Gateway.

2.3 Typical Wireless Network Operations

The 3G and 4G mobile wireless networks are quite complex, with various mobile device and network operations combining to support high data speeds and ever-improving quality of service. Figure 3 provides examples of such operations of mobile devices and the network for LTE.

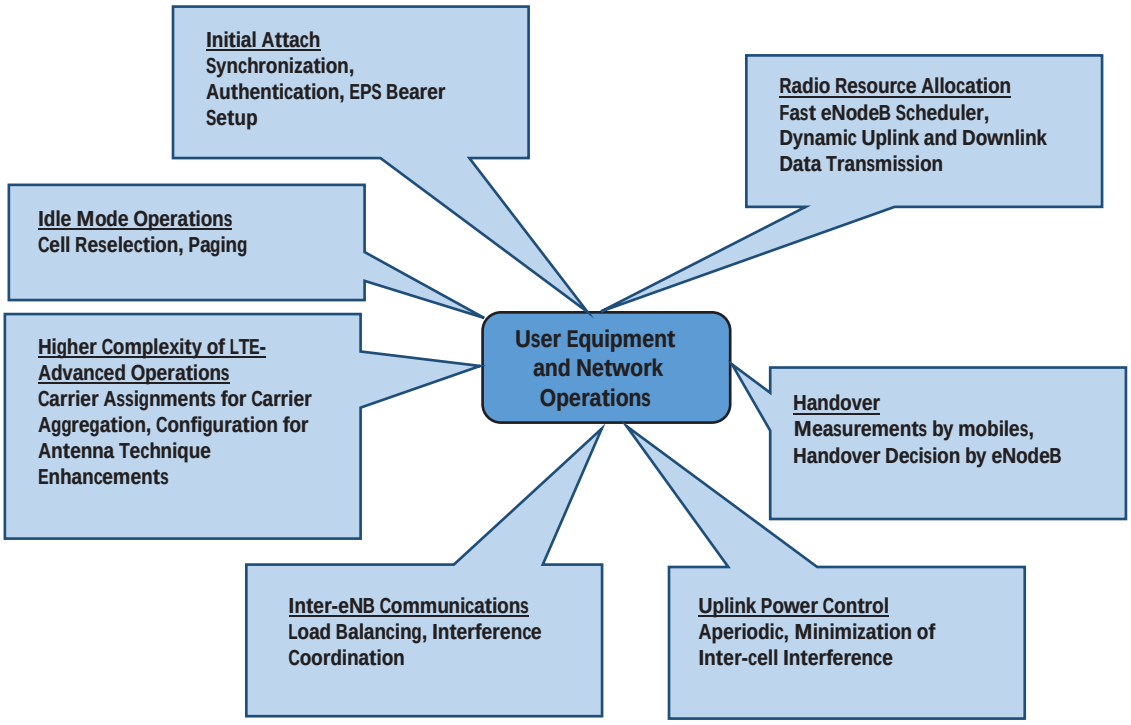

Figure 3. Operations of the Wireless Mobile Device and the Network

The mobile device carries out an initial attach procedure after power-up. During the attach procedure, the mobile achieves downlink and uplink synchronization with the eNodeB. The mobile and the network authenticate each other, and security is established. A default Evolved Packet System bearer with best-effort service is established toward a default packet data network to carry information without any

[2] End-to-end QoS is controlled at the EPS bearer level in LTE. Hence, if two applications need two different levels of QoS, two different EPS bearers with distinct QoS characteristics are needed. Furthermore, two applications with same QoS needs can be placed onto the same EPS bearer.

guaranteed data rate but with the target delay of 300 ms between the mobile and the Packet Data Network Gateway. The mobile is typically allocated an IP address during the default Evolved Packet System bearer setup.

Active mobiles have one or more Evolved Packet System bearers, and, the eNodeB scheduler dynamically allocates radio resources to the mobile for the downlink data transmission and the uplink data transmission. The eNodeB scheduler executes as fast as every millisecond to adapt to the radio channel conditions and to modify the allocated downlink and uplink resources.

The serving eNodeB configures the active mobile with measurements of neighboring cells that can be on the same carrier frequency as the serving cell or a different carrier frequency, or a different radio access technology (e.g., UMTS). The mobile device provides measurement reports when configured measurement events occur. The serving eNodeB makes a handover decision (if appropriate) and works with the target eNodeB to obtain resources for the mobile. Handover may occur without the movement of a user if the handover would balance traffic between eNodeBs.

In addition to allocating spectrum resources to the mobile, the eNodeB also controls the transmit power of the mobile by sending power control commands. Power control in LTE may be implemented as aperiodic and multiple power step-up and step-down sizes can be used. Power control helps minimize inter-cell interference in the uplink.

The eNodeB may communicate with the neighboring eNodeBs to carry out load balancing and to coordinate interference. Minimizing interference improves the achievable user throughput and cell throughput. Scheduling provides a compromise between fairness in serving all users and throughput for the overall network.

Complexity of the LTE network increases further with LTE-Advanced. The eNodeB scheduler needs to decide when to use multiple carrier frequencies simultaneously for a given mobile to improve throughput as part of the carrier aggregation feature of LTE-Advanced. More antenna technique enhancements are available in LTE-Advanced compared to LTE, and, the eNodeB dynamically needs to determine the type and configuration of the multiple antenna technique.

In the absence of data activity for a configurable time period, the eNodeB asks the mobile to enter the idle mode. The network needs to keep track of mobiles in the idle mode so that the network can page the mobile in the correct geographic region for incoming voice or data traffic. Even though the mobile in the idle mode does not consume any radio resources, it performs cell reselection to observe the strongest cell so that it is in the best possible cell when it needs to exit the idle mode to do some activity such as signaling exchange or data transfer.

3. Characteristics of Mobile Wireless Networks and Differences Between Wireless Networks and Wireline Networks

In Section 3.1, the characteristics of mobile wireless networks are discussed in detail. These characteristics dictate the complexity of network management and the need for flexibility for wireless providers to respond to changing circumstances within the network. Section 3.2 describes the significant differences between wireless and wireline network architectures that warrant differences in how mobile wireless networks are managed.

3.1 Characteristics of Mobile Wireless Networks and Resulting Implications

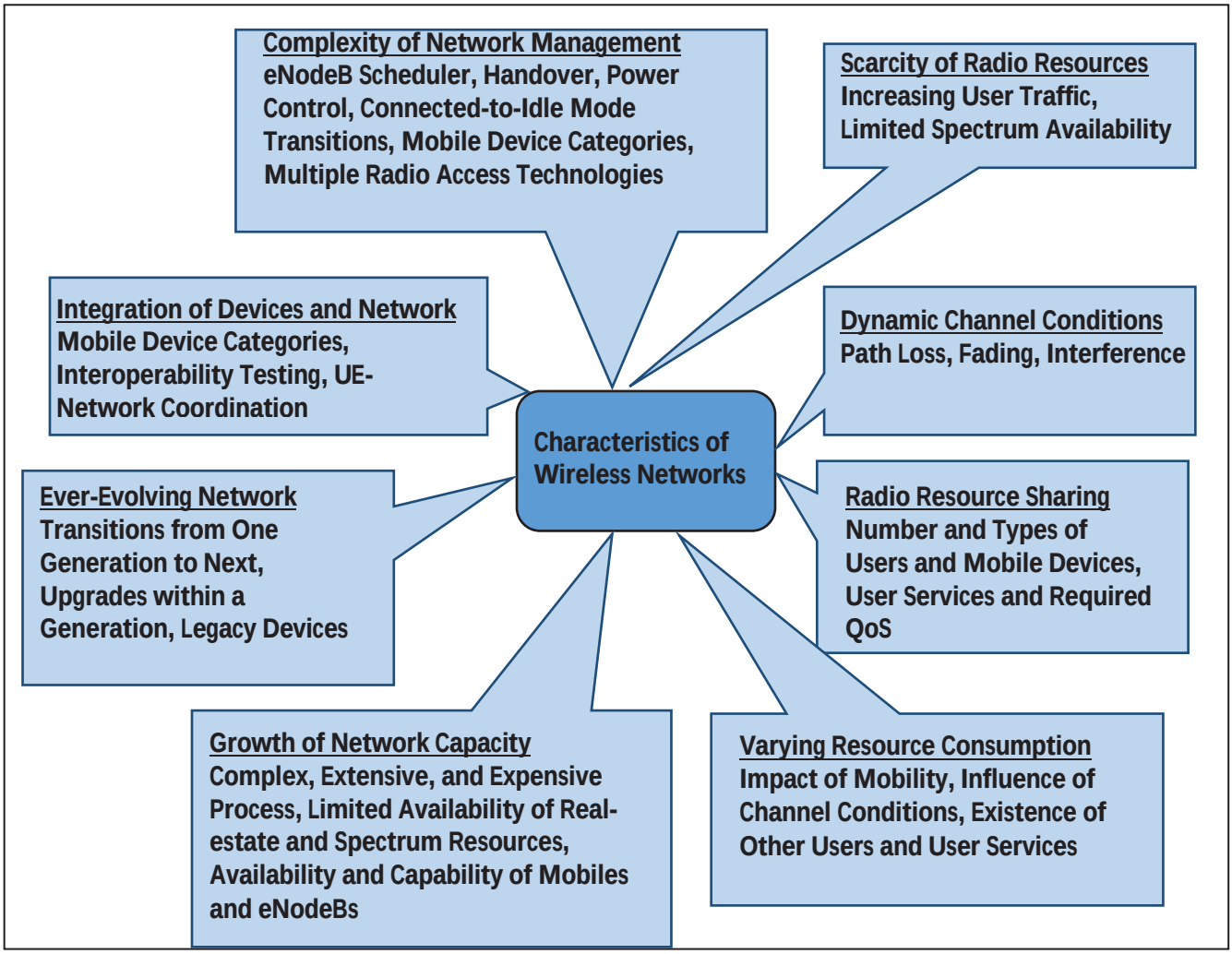

Figure 4. Characteristics of Mobile Wireless Networks

Figure 4 summarizes the characteristics of mobile wireless networks. These characteristics as a whole pose significant challenges to mobile wireless networks and make the application of prescriptive net neutrality principles to mobile wireless networks practically infeasible. In particular, determination of any reliable thresholds to quantify user or network performance is impossible. Furthermore, service providers need a high degree of flexibility to efficiently and effectively manage precious radio resources to ensure the best possible service experience for all subscribers.

Scarcity of Radio Resources. The popularity of the Internet and IP-based services such as video streaming have contributed to the explosion in the amount of data traffic traveling through the mobile broadband network. 4G services such as LTE bring with them higher data speed and greater video quality. The result has been more intensive use of 4G devices for bandwidth-heavy services, such as streaming video. Globally, in 2013 a 4G connection generated 14.5 times more traffic on average than a

non-4G connection.[3] Although 4G connections represent only 2.9 percent of worldwide mobile connections today, they already account for 30 percent of mobile data traffic worldwide.[4] In the United States, the average 4G smartphone generated 1,739 MB of traffic per month in 2013, compared to 906 MB for non-4G smartphones.[5] Cisco estimates that "In the United States, mobile data traffic by 2018 will be equivalent to 383xthe volume of U.S. mobile traffic ten years earlier (in 2008)."[6] However, spectrum does not become available with the same growth rate as data traffic. Mobile broadband operators are constrained by the amount of spectrum available and the growth rate of new spectrum availability will not keep up with constant increases in user demand. This is exacerbated by the rapid rate of data intensive applications, now enabled by mass adoption of screen based smartphones and tablets that encourage use of pictures, graphics and video, and hence drive data demand as well as driving requirements for lower latency (real time response). Scarcity of radio resources, such as spectrum, necessitates efficient management of aggregate radio resources that needs to strike a balance among numerous competing factors such as the number of active users, target QoS of user services, and prevailing radio channel conditions.

Radio Resource Sharing. Limited radio resources must be shared among the active users in a given geographic area. Basically all of the channel capacity is divided among the various users and the speed for every user will go down as more users are added. A small number of very heavy data users using apps that are extremely data intensive can have a disproportionate impact on a large number of users. The eNodeB scheduler, as often as every millisecond, needs to consider a number of factors such as the number of active user devices, capabilities of these devices, capabilities of the eNodeB, prevailing channel conditions of different devices on the network, and target QoS of different services to determine the amount of radio resources for individual users. Even if best-effort service were the goal for all users, these users would typically experience different data rates as the eNodeB scheduler would try to improve overall network throughput and overall user throughput.

Dynamic Channel Conditions. For a given level of service quality, the required amount of radio resources is a function of the channel conditions, and the channel conditions not only vary over time, but also as a function of distance from the serving cell. The signal-to-interference plus noise ratio (SINR) directly influences the required radio resources. SINR is influenced by a variety of factors such as the propagation-based signal attenuation, the severity of fading (e.g., shadow fading and Rayleigh fading), and the amount of interference. Furthermore, the channel conditions hardly remain static. The channel conditions change due to factors such as user mobility. Network operators need maximum flexibility to manage radio resources to quickly adapt to changing channel conditions. Even to preserve a given data rate, the user may need 36 times more radio resources when the channel conditions degrade.[7] For

---

[3] Cisco, Cisco Visual Networking Index, Global Mobile Data Traffic Forecast Update, 2013-2018 at 2 (Feb. 5, 2014) ("Cisco Feb. 2014 VNI Report"), available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf.

[4] *Id.*

[5] Cisco, VNI Mobile Forecast Highlights, 2013-2018 at "United States – Accelerating Network Speeds" ("Cisco VNI Highlights"), *at* http://www.cisco.com/assets/sol/sp/vni/forecast_highlights_mobile/index.html#~Country (last visited June 10, 2014).

[6] *See* Cisco VNI Highlights at "United States – 2018 Forecast Highlights."

[7] To quantify downlink channel conditions, the LTE standard has defined Channel Quality Indicator (CQI) that is a measure of achievable spectrum efficiency. CQI=1 corresponds to poor channel conditions, while CQI=15 corresponds to excellent channel conditions. The efficiency of transmission decreases from 5.5547 bits to 0.1523

example, a far-away user may require more coding (effectively more redundancy, meaning a higher real radio data rate to support the same effective data rate) and more retries (faulty packets with too high an error rate to be properly decoded are resent). Thus not all users are the same, even though their perceived data rates (the data rate the end user observes) appear the same. There are no definable metrics that could 'fairly' assess the achieved data rate. It takes the network effectively more network air interface resources (radio capacity) to serve such far-away (poor radio channel) customers. There is no such analogous situation for wired or fiber optic networks, because the channel quality conditions do not vary by such a large ratio, nor are the channel conditions so variable over time or space.

Varying Resource Consumption. Users in different channel conditions and using different services consume different amounts of resources. Even for the fixed throughput, different users would consume different amounts of radio resources depending upon the device-specific channel conditions. For a given channel condition, different services such as email and a VoIP call would consume different amounts of resources. It is nearly impossible to determine the exact amount of radio resources for a given user due to the highly dynamic nature of mobile wireless networks.

Challenges of Network Capacity Additions. Mobile broadband providers invest heavily to increase network capacity and keep up with rising user traffic and user expectations. Capacity can be increased by adding more spectrum (more different bands or more channels within the existing band(s)), deploying capacity-enhancing features such as advanced antenna techniques, and adding more cells (either by deploying 'split' macro cells or small cells) via cell-splitting techniques to gain more capacity via more 'frequency reuse.' In general, many of these techniques are quite expensive and take a long time from the concept to full commercial realization. Also, many of these radio capacity enhancing techniques have practical limitations. Deploying multiple bands requires replacing the users' handsets, and the costs rise as the devices are more complex to serve multiple bands. Base station cell splitting techniques cannot be implemented indefinitely because co-channel interference levels rise as the cells get smaller. Advanced antenna techniques require larger antenna arrays.

Thus, as noted above, mobile wireless broadband providers cannot simply build their way out of capacity constraints but instead are dependent on government allocation of spectrum resources and must purchase rights to use these resources at auction. Purchasing spectrum resources and implementing other capacity-increasing techniques can be quite expensive. Adding macro cells poses an additional challenge of finding real estate. To exploit the full potential of the standard, the user equipment and the eNodeB need to have compatible capabilities. It may take years before the commercial incarnations of user equipment and the eNodeB are coordinated and can deliver the target theoretical peak performance aimed by the standard. The intricacies of capacity growth along with ever-rising user traffic imply that efficient utilization of the existing radio resources is absolutely critical to the user experience and the network efficiency.

Ever-Evolving Network. As mentioned in Section 2, the mobile broadband service provider's network keeps changing to adapt to the newer generations of cellular standards and multiple revisions within a given generation of the cellular standard. The network has to manage the user equipment (UE) across various generations and revisions. As the newer standard emerges, the older standard does not

---

bits for a given modulation symbol, leading to 5.5547/0.1523=36.4 more resources under the poor channel conditions to preserve a given data rate in poor and excellent channel conditions.

disappear immediately. Even the first-generation analog standard in the U.S., AMPS, survived for more than two decades! There are wide variations in achievable network performance and user-experienced QoS change among generations and even revisions within a generation. For example, a user may experience data rates of tens of Mbps (megabits per second) in an LTE network, but this speed could go down to hundreds of kbps (kilobits per second) when the user enters a UMTS network. Such wide disparity of the achievable performance makes it difficult to quantify even the minimum level of QoS or any metric (used for assessing performance and network neutrality) that relies on the apparent user experience.

Integration of Devices and Network. The user equipment and the network need to be tightly integrated to ensure satisfactory user experience. The standards typically define multiple categories of user equipment with different capabilities. Common ground needs to be found between a given category of user equipment and the eNodeB. In LTE, the network learns about the capabilities of the user equipment during the initial attach procedure and properly configures the equipment to ensure seamless communications between the device and the network. The network often works with user equipment of differing capabilities. Hence, even when two devices have identical channel conditions and identical allocation of radio resources, they could experience widely different throughput depending upon their capabilities as well as the proprietary aspects of the devices, such as antenna design. Extensive integration testing is carried out to ensure proper operations of user equipment and the eNodeB and error-free interactions between the device and the network. Tight integration between the user equipment and the network (e.g., eNodeB, Evolved Packet Core, and IMS) plays an important role in ensuring good user experience. Again, no 'fair' metrics could be defined to account for such differences in performance.

Complexity of Network Management. The network management in modern mobile wireless networks is extremely complex. Numerous interactions among the user equipment, the eNodeB, the Mobile Management Entity, the Serving Gateway, the Packet Data Network Gateway, the IMS network, and the Policy and Charging Control network occur to provide seamless communications experience and end-to-end QoS to the user. As mentioned in Section 2, the eNodeB scheduler allocates radio resources for the downlink and the uplink data transfer to achieve target QoS levels for the established Evolved Packet System bearers. The eNodeB executes a handover algorithm to choose the best possible serving cell for a user. The eNodeB also manages uplink power control commands to the mobiles to minimize inter-cell interference. The user equipment would be allowed to transmit more power if its uplink channel conditions are poor and/or its uplink throughput requirements are high. The eNodeB and the Mobile Management Entity manage connected-to-idle transitions for the user equipment. The network management must consider different capabilities of different mobile device categories to optimize the experience for the user. Ensuring seamless mobility across different radio access technologies is a non-trivial task. The network needs to configure the user equipment with suitable measurements and needs to connect radio networks supporting different radio access technologies. Integration testing within the network is also required to verify error-free coordination across radio access technologies. Nevertheless, this cross-layer optimization of the overall network is important for overall system performance and continues to be a promising area for further improving overall network performance.

3.2 Differences Between Wireline Networks and Mobile Wireless Networks

Any proposals to extend network neutrality principles conceived in a wireline context to mobile operations must contend with the vastly different technical challenges of these two types of communication networks. This section provides an overview of the differences in technical challenges between wireline and wireless systems as they relate to network neutrality regulation.

Wireless channels are quite different from wireline channels. First, the bandwidth for a wireless service provider might be on the order of 10s of MHz (~10^7 Hz) (5-30 MHz), but a fiber optic system could be 10s of GHz (~10^{10} Hz). The difference represents at least a one thousand-fold difference and in many cases is much greater in total bandwidth. The number of users or data rates that can be accommodated is directly proportional to the total bandwidth (and, in wireless systems, is also affected by the relative dispersion of the users within particular cells). Although 3G and 4G technologies can enable multi-megabit per second wireless transfer rates (assuming adequate spectrum resources), wireless systems will never have the bandwidth of wireline systems. A wireline network can exploit advances in optical fiber technologies to achieve extremely high bandwidth exceeding thousands of Gbps (gigabits per second). In contrast, the limited amount of radio spectrum in mobile wireless networks puts a severe constraint on the achievable data rates on a wireless link. Additionally, the wireline network is very consistent with respect to capacity capabilities of the channel over time (no fading) and space (low loss per distance of fiber). The wireline network engineer knows precisely how much bandwidth is available in a single fiber optic strand and (other than losses over distance) will have a near-constant understanding of the performance of the transport layer. In contrast, wireless networks are faced with ever-changing radio environments. Temporal issues such as multipath, clutter, blockage, channel fading, and extraneous interference will result in changes in the performance of the network and the quality of service experienced by subscribers. Also, the quality of the radio channel necessarily degrades rapidly as a function of distance from the serving cell. Without extensive management (and the inherent compensation mechanisms used within the radio air interface: variable rate coding, variable modulation, retry, etc.) of the wireless network to account for these transport layer issues, customers would not receive the types of services and data rates that they expect.

Moreover, a “build more infrastructure” approach is much less of a solution to capacity issues in wireless systems than in wireline systems for a number of reasons. First, spectrum constraints place outside limits that simply do not exist in wireline. Overall aggregate wireline bandwidth can be expanded infinitely by adding more cables or fibers, or by technology upgrades. Wireless bandwidth is ultimately constrained by fundamental performance limits, available spectrum and interference. Second, mobility and propagation issues combine to create much greater variability in the channel as compared to wireline channels. Third, mobility and propagation issues combine to create much greater variability in wireless traffic—the spread between peak and average traffic levels is typically much wider for wireless than wireline—which makes it infeasible to design networks to meet anything approaching peak demands. Fourth, issues unique to wireless networks are associated with deploying more capacity. Wireless carriers continue to spend billions of dollars annually on infrastructure upgrades, but they will continue to face severe capacity constraints, particularly with demand growing far faster than anticipated and faster than new bands can be added.

In wireline systems, in contrast, capacity improvements without the large expense of laying new fiber have been made possible through better technology at the fiber ends. Such technology options simply

are unavailable for wireless systems, and dynamic prioritization and other management techniques are and will remain essential. While wireless network providers have taken efforts to use their spectrum resources more efficiency, such as by using small cell technology, as explained above wireless operators simply cannot "build out" of capacity constraints to the same extent as their wireline counterparts. In the 30 year history of commercial mobile networks, wireless providers have moved from analog (1G) to digital (2G) to 3G and now 4G services. However, each radio interface change requires substantial time and investment to bring about the gains in efficiencies expected from the more robust standards. Each base station must be updated via software and/or hardware to accommodate the changes in the air interface. All of the existing mobile devices in the network must be replaced to provide the full benefits to spectrum efficiency that the new radio standards allow. In contrast, wireline networks are able to upgrade solely at the edge of their networks to help gain efficiencies and do not require the extensive costs associated with wireless network technology migration to provide capacity gains. Fiber also presents extensive capacity availability throughout the network that has not yet been tapped for use, but is readily available for carrying traffic with updates to the technology at the fiber ends. Not only is the bandwidth of the wireless channel severely constrained compared to wireline channels, the reliability of the wireless channel is well below that of a wireline channel. The reliability issue is due to a number of factors, such as blockage of the radio signal (called shadowing), echoes or multipath of the signal, thermal noise, and, more importantly, interference. These impairments to the channel create substantial additional complexity and variability. Planning and operating a wireless deployment to ensure Quality of Service (QoS) and coverage is extraordinarily difficult because these impairments are random and unpredictable.

Interference is often the most important of these impairments, and, by its very nature, is constantly changing between and within cells. Interference occurs when multiple signals share the same spectrum. These signals are typically associated with the same service provider but are sometimes due to another service provider using the same or adjacent spectrum bands. Interference limits capacity in a wireless system on a dynamic basis, varying by location and from one millisecond to the next, and this problem has no counterpart in wireline systems.

Deployment and maintenance of wireline systems is less dynamic than wireless systems. Although wireline electronics and services continue to evolve, the advent of fiber has brought relative stability and efficiency to the wireline network architecture. In contrast, only change is constant in wireless standards and networks. As a result, network management practices must constantly evolve to address new architectures, new technologies, new standards, and new wireless applications with new performance needs.

These various features of mobile wireless networks make them much different than wireline networks. Table 1, below, summarizes the differences between wireless and wireline networks.

Table 1. Summary of Differences Between Wireless and Wireline Networks

Characteristic	Wireline	Wireless
Communications Channel	Relatively clean with signal Regeneration	Impaired with noise, interference, multipath, and blockage
Bandwidth	No spectrum limitations	Severe Spectrum limitations
Mobility	None	Constant, complex, often unpredictable, and often consuming extensive resources
Power	No need to manage power/battery life in wireline network for end user devices.	Limited power/battery on user device that must be accommodated through network management
Security	A lesser concern due to the physical path between the provider and the user (buried or on aerial infrastructure).	A greater challenge due to the possibility of tracking a user and variety of interfaces
Response to Increased Traffic Demand (i.e., the Capacity Problem)	Capacity increases may be feasible, although soaring demand and increasing congestion issues may call for additional pricing, bandwidth limitations, and prioritization mechanisms	Primarily managed dynamically through prioritization, scheduling, and power allocation
Network Complexity	Relatively simple	Extremely complex
Network Stability, Deployment, and Maintenance	Comparatively stable platform and systems, although high growth in demand and new applications are issues	Extremely dynamic platforms and systems; Deployment and maintenance require constantly dealing with real estate acquisition and zoning issues; Planning and maintenance are more difficult, and continuous maintenance and frequent resetting of network parameters is required; Infrastructure changes to address localized capacity issues can have ripple effects through adjacent cells

Characteristic	Wireline	Wireless
Quality of Service	Easier to implement due to availability of higher capacity and predictability of resource requirements	Quite difficult to implement due to variable capacity, unpredictability of resource requirements, and existence of proprietary mechanisms; Industry moving toward IMS and PCC

4. Challenges of Implementing the FCC's Proposed 2014 Rules on Net Neutrality to Mobile Wireless Networks

The NPRM seeks feedback on the *transparency rule*, the *no-blocking rule*, and the *anti-discrimination/commercial reasonableness rule* in the context of mobile broadband service providers. The NPRM proposes to apply the transparency rule to both fixed and mobile broadband wireless access. Regarding the no-blocking rule, the NPRM proposes to treat mobile and fixed broadband services differently. Furthermore, just as the FCC chose not to apply the 2010 unreasonable discrimination rule to mobile broadband service, the 2014 NPRM tentatively concludes that the replacement rule – or "anti-discrimination/commercial reasonableness" rule – would not be applicable to mobile broadband. Section 4.1 discusses the challenges of applying the enhanced transparency rule to mobile wireless networks. Section 4.2 describes the problems encountered while applying the enhanced no-blocking rule to mobile wireless networks. Section 4.3 briefly explains why the NPRM's view of not applying the unreasonable discrimination rule and the "anti-discrimination/commercial reasonableness" rule to mobile wireless networks is the correct approach. Extensions of the transparency and no blocking rules beyond those adopted in 2010 would be unwieldy and over-inclusive. Application of an anti-discrimination/commercial reasonableness rule to mobile broadband providers would hamper their ability to innovate, optimize, differentiate, and deliver high quality products and services.

4.1 Transparency Rule and Mobile Wireless Networks

The 2014 NPRM seeks comment on expansions of the transparency rule that would require mobile service providers to disclose information in several categories, including *network management practices*, *performance*, *congestion specifics (e.g., speed and packet loss)*, *peak load management*, and *parameters of default or best-effort service*. However, as explained below, for mobile providers there are numerous technical and practical problems in meeting these proposed expanded disclosure requirements that make implementation of any enhanced transparency rule problematic, resulting in increased costs, less responsive service due to limitations on network management and would not provide consumers with relevant or useful information.

4.1.1 Network Management Practices

In a typical wireline network, the only variable is the amount of traffic on a given link – all other things such as capacity, etc. are typically static. This makes management of the traffic relatively straightforward using standard queuing techniques (e.g. Weighted Fair Queuing) to ensure all customers receive a fair share of the available bandwidth during congestion caused by a small number of users.

With wireless networks, there are many variables that are all changing simultaneously – signal strength and interference affect capacity, orientation of antenna affects throughput, and obstacles can dynamically interrupt data, among other things. Using just standard wireline techniques would not work well in this environment, and as described below, there are many methods used to make the network function well. During times of congestion, heavy users may have to be treated differently based on multiple variables to ensure proper throughput for all users. This is something accounted for in the standards as well as in most network management practices, and requiring all these technically-driven capabilities to be suspended simply to ensure a "neutral" network can have significant negative consequences.

eNode B Base Station. The *network management practices* in mobile wireless networks are extremely complex and consist of numerous mechanisms that are distributed among various nodes in the network architecture illustrated in Section 2. The achievable radio network performance and user experience are influenced heavily by these network management mechanisms. Sections 2 and 3 identified several network management mechanisms implemented by the eNodeB such as the scheduling algorithm for downlink and uplink resource allocation, the handover algorithm, the load balancing algorithm, handling of the connected mode-idle mode transitions, adaptation to the changing channel conditions, power control, and interfere coordination. Although the standard defines auxiliary tools such as (i) measurement reporting by user equipment and (ii) inter-eNodeB signaling exchange via the standardized X2 interface, these network management mechanisms are proprietary to the infrastructure vendors. Infrastructure vendors differentiate their products based on abilities of these mechanisms. Hence, a mandate to fully disclose these mechanisms would discourage innovations, violate intellectual property rights, and harm both competition and consumers.

Core Network. Just like the network mechanisms implemented at the eNodeB, the network management mechanisms implemented in the Evolved Packet Core and the auxiliary networks of IMS and the Policy and Charging Control networks[8] could provide a competitive edge and serve to differentiate service providers. The load balancing among the Mobile Management Entities and management of idle mode mobile devices are examples of network management in the Evolved Packet Core that are vendor-proprietary. The service provider may have a specific way of providing a certain level of QoS for a given service by configuring the IMS and the Policy and Charging Control networks (e.g., certain target data rates and certain latency targets). Furthermore, the routers that carry signaling and user traffic between the eNodeB and the Evolved Packet Core and within the Evolved Packet Core may be configured by the implementation-specific network management framework.

Service providers need maximum flexibility in the network management of mobile wireless networks to make the best use of the *scarce radio spectrum* in the presence of the exponentially rising data traffic. For example, [Neel_MobileDataTraffic] reports that the mobile data traffic is expected to grow by a factor of 450 from 2005 to 2015. Furthermore, *scarcity of the radio spectrum* is clearly evident in Exhibit 11 of [Deloitte_SpectrumShortage], where the FCC estimates that the U.S. would experience a spectrum

[8] While we have mentioned examples of major network management mechanisms, we note that these are not the only mechanisms that exist. Numerous other algorithms that manage radio resources, core network resources, transport network resources, IMS and PCC resources exist. For example, some mechanisms to configure the operations of the radio channels and to coordinate resource utilization between macro cells and small cells would be needed.

deficit of 275 MHz relative to the demand in 2014. Due to the scarcity of precious spectrum, innovative, high-performance, and ever-evolving network management mechanisms are absolutely essential to the overall network performance and user experience. Flexibility in tuning and adapting the network management mechanisms to fast-paced technology evolution, implementation of new features and the uncertainty of the requirements of emerging applications or services require that the network management mechanisms in mobile wireless networks should not be subject to any disclosure requirements. In the fast-paced evolution of wireless standards, multiple revisions exist, as discussed above. Even for a given revision of the standard, the user equipment and the network vendors have multiple software releases to update. Revelations of the network management mechanisms in the eNodeB, the Evolved Packet Core, and the IMS and Policy and Charging Control networks would ultimately harm consumers due to a reduced rate of innovation resulting from adherence to counter-productive implementation of the transparency rule.

The situation is too complex to summarize with a small set of easily defined comparative metrics, and expansion to a more complex, detailed set of base station and/or network performance metrics would violate the confidential nature of the network providers' proprietary technical optimization choices. This would severely impact the pace of innovations in the area of network management.

4.1.2 User Experience and Network Performance

Performance is extremely difficult to estimate reliably for mobile wireless networks, because there are numerous factors that influence the achievable network performance perceived and user performance (user experience). To complicate the performance estimation further, many of these factors are dynamic. Also, many of these factors are application specific, in that they are more important for some applications than for others. As discussed in Section 3, the wireless network is characterized by a variety of factors including dynamic channel conditions, varying number of active users, differing QoS requirements for the services of active users, the available amount of spectrum, user mobility, the capabilities of user equipment and the eNodeB, the types of applications considered, and the generation and the revision within the generation of the wireless standards. All these factors together determine the network and user performance at a given instant. Furthermore, this performance would change from one instant to another as one or more factors change. Reporting average performance over a particular period of time may not make sense and could be misleading to consumers.

The achievable user throughput is a function of the signal-to-interference plus noise ratio, which is influenced by the radio channel conditions that reflect propagation-based path loss, type and severity of signal fading, and amount of interference. Signal-to-interference plus noise ratio, in conjunction with other measurements[9], also dictates the configuration of the advanced antenna techniques that can be used for a given mobile device at a given instant.

The number of active users and the specific QoS requirements of the services of these users determine how the available radio resources of the network are distributed. For example, in LTE one active user could get up to around 75 Mbps in a case of excellent channel conditions that are conducive to the use of spatial multiplexing technique. In contrast, poor channel conditions resulting from a weak signal (e.g., due to fading and large propagation path loss because the user is far away from the serving base

[9] Examples of these measurements include rank indication (RI) and precoding matrix indicator (PMI).

station) and strong interference may be able to support only about 1 Mbps.[10] If there are multiple active users with guaranteed bit rate (GBR) requirements, other users involved in non-GBR services such as email and web browsing will experience much lower average throughput. As noted above, the numerous variables inherent to a wireless network may mean that during times of congestion, heavy users may have to be treated differently to ensure proper throughput for all users. Should the FCC mandate the suspension of such network management practices in the name of "neutrality," significant negative consequences could result.

The throughput experienced by a user would also depend on the service (or application) being received by the user. A seemingly low instantaneous data rate of about 300 kbps would be more than adequate for a VoIP call, while a much higher instantaneous data rate (e.g., few Mbps) would be needed for video streaming or a file download to provide satisfactory user experience. In the absence of the context of the specific service or application and their related QoS requirement, a given value of data rate, or any other metric, such as latency, is not a reliable indicator of the user performance or experience.

The available amount of spectrum directly affects the achievable performance. If a service provider has a 10 MHz LTE channel in one market but only a 5 MHz LTE channel in another market, the achievable throughput can easily differ by a factor of more than two. The larger the channel bandwidth, the higher the achievable throughput. Frequency selective scheduling could provide larger gains in case of larger channel bandwidths.

The impact of user mobility on achievable performance is also significant. In general, a higher velocity of the user equipment results in a larger Doppler shift and typically implies frequent and more severe short-term fades. In contrast, a slowly-moving device (e.g., pedestrian speed) has a smaller Doppler shift and experiences fewer varying signal fades, but the time period for the fade may be much longer and more impactful, as the user may remain within a performance null area for a longer period of time. The signal-to-interference plus noise ratio required to achieve a target throughput (or error rate) can vary significantly due to the impact of the user mobility and the distance from the serving cell.

The network performance and the user performance are affected by the capabilities of the user equipment and the eNodeB. All the eNodeBs do not have support for all the configurations defined by the standard. For example, Release 8 LTE supports parallel transmission of data from four antennas. However, commercial Release 8 LTE deployments typically have two transmit antennas for parallel data transmission. Similarly, not all devices have these same capabilities. Five categories of user equipment are defined for release 8 LTE, and, Category 3 devices are widely used in current commercial deployments in the U.S. and around the globe. Category 3 devices receive signals on two antennas. Furthermore, Category 3 equipment supports QPSK and 16-QAM modulation schemes to transmit data in the uplink. In contrast, Category 5 devices have four antennas to receive the downlink signals and support QPSK, 16-QAM, and 64-QAM to transmit data in the uplink. Such differences in equipment categories are a key reason why the downlink peak data rate is around 300 Mbps for a Category 5 device and around 100 Mbps for a Category 3 device and the uplink peak data rate is around 75 Mbps for a Category 5 device and around 50 Mbps for a Category 3 device [3GPP_TS36.306].

[10] These calculations assume that a (2x2) MIMO (Multiple Input Multiple Output) technique is used when the channel conditions are excellent (i.e., with CQI=15) and a non-MIMO technique at CQI=1 is used.

The generation and the revision within the generation of the wireless standards also have a significant influence on the achievable performance. For example, Release 8 LTE supports theoretical peak data rates of 300 Mbps downlink and 75 Mbps uplink. In contrast, Release 10 LTE-Advanced supports a theoretical peak data rate of 3 Gbps downlink and 1.5 Gbps uplink. Again, defining one set of reasonable metrics is impossible in the face of different generations of handsets.

In summary, mobile broadband providers are committed to complying with the existing transparency rule but, recognizing that the specifics of network and handset performance will vary constantly, the granularity contemplated by the proposed expanded transparency rule would be infeasible. There are simply too many factors (most of which are highly dynamic or variable) that influence network and user performance, making it impractical to predict, guarantee, and/or verify user performance in the context of the expanded transparency rule.

4.1.3 Congestion

The 2014 NPRM envisions disclosure of congestion-related statistics such as speed (i.e., throughput) and packet loss. In the context of mobile wireless networks, such disclosure requirements are unhelpful for several reasons explained below.

First, the dynamic nature of mobile wireless networks leads to wide variations of throughput as discussed in Section 4.1.2 above. For data-centric systems, packet loss is reflected in the overall throughput. Furthermore, some throughput degradation and/or packet loss may not be due to congestion at all; it may simply be due to changes in channel conditions, user mobility, and/or service change. The packet loss rate may not really reflect any congestion.

Separating congestion issues from non-congestion issues through analysis and storing and maintaining such data for the sole purpose of compliance with the transparency rule would need significant investment of resources without any tangible benefit to consumers. As mentioned in Section 3, the eNodeB scheduler operates as fast as every millisecond, and the number of active users can change within a few seconds due to transitions between the idle and connected modes. So-called congestion can widely fluctuate in matter of few seconds. Network optimization processes aimed at addressing congestion can respond to these temporary congestion issues just as quickly. As a result, real-time disclosure of network congestion would be problematic to implement and confusing for consumers, as the network is constantly responding to ever-fluctuating levels of traffic. Section 2 illustrated the LTE network architecture. The complexity of the overall architecture means that network upgrades due to revisions or new feature implementations would almost certainly need extensive troubleshooting efforts which would likely lead to temporary congestion issues. Engineers seeking to comply with such a rule would face an unnecessary burden that could delay solutions to real problems, and cause harm to the subscribers instead of helping users (which is the real goal of the transparency rule). The key to keeping overall network capacity high is to adapt to the traffic over time using sophisticated and proprietary scheduling algorithms. Network neutrality rules which place strict demands on traffic handling and may restrict or prevent certain schedule techniques, would require infrastructure and capacity to be over-engineered to handle otherwise manageable peaks, and hence result in higher costs for consumers.

Since network optimization is an ongoing and iterative process, it is quite likely that some congestion issues that are reported will have been remedied long before they could be incorporated into any required disclosure. Mobile wireless networks have numerous challenges on the radio channel, and,

service providers need to invest heavily in ongoing technology upgrades, network RF planning, design, and optimization activities. The technology for wireless is changing very quickly and will continue to change quickly for the foreseeable future. Service providers will have a learning curve to understand how to deploy this technology and realistically, regulatory policy will (and perhaps must) adjust more slowly to these technology developments. Undue regulation at this point will stifle technology deployment that could increase bandwidth availability and lower costs for the consumer.

4.1.4 Peak Load Management

There are many legitimate reasons why a wireless network operator needs to manage data traffic on its network. In such cases, reporting of such peak load management would have little benefit. Recall from Section 2 that the Policy and Charging Control network works with the Evolved Packet Core to ensure suitable QoS. Such interworking between the Policy and Charging Control network and the Evolved Packet Core means that peak load management of traffic may be carried out such that the QoS for a given Evolved Packet System bearer is met. Each Evolved Packet System bearer has QoS parameters such as the maximum data rate. If the incoming data rate exceeds the subscribed data rate, the Packet Data Network Gateway manages the data to meet the data rate constraint toward the Serving Gateway. Such peak load management is carried out as part of the 3GPP standard's QoS characteristics [3GPP_TS23.203]. The network needs to manage the traffic so that all users can satisfactorily receive services instead of just few users consuming disproportionate amounts of resources. If excessive amount of traffic is received at the eNodeB, the eNodeB may have to buffer the packets, delaying the packets and potentially causing packet loss if the device-specific buffer overflows. Hence, even for the user with higher data rate needs, suitable peak load management is needed.

The resources in mobile wireless networks are scarce, and these scarce resources must be shared among numerous users. If suitable optimization is not carried out, some applications could flood the network with excessive amounts of data traffic (and even signaling traffic), causing degradation to many users in the network. In general, higher data rates result in the consumption of more resources, and, concentration of radio resources among only few subscribers would be unfair to other users. Hence, network optimization could be viewed as a network management mechanism to provide some level of fairness among the uses and users of radio resources. Network operators need the flexibility of such legitimate management to strike a balance among fairness, network performance, and aggregate user performance. The network management in mobile wireless networks must have at its disposal all means, including optimization, to safeguard the interests of all subscribers and to provide the best possible experience to all subscribers instead of just a select few subscribers. This balancing is dynamic and load dependent, so again, one set of uniform metrics cannot meaningfully capture overall performance.

4.1.5 Parameters of Best-Effort Service

The NPRM asks if any parameters can be specified to quantify the best-effort service. Numerical parameter settings that quantify the best-effort service and that are reliable are difficult to guarantee in mobile wireless networks. First of all, commercial mobile wireless networks have a mix of radio access technologies, mobile devices with different capabilities, and eNodeBs with different capabilities. Hence, the achievable performance varies depending upon the specific combination of the technology, the mobile devices, and the eNodeB for a given channel condition.

Even within the narrow scope of a given standard such as LTE and ignoring differences among technologies, compliance with the NPRM-proposed transparency rule in the form of target parameters would be quite challenging. LTE defines nine levels of QoS in terms of QoS Class Indicators (QCIs) [3GPP_TS23.203]. A QoS Class Indicator (QCI) specifies the QoS class. Defining different data rates for these services offers operators additional flexibility. An operator could also define proprietary QCIs. For example, QCI = 1 is suitable for applications such as VoIP. Its priority is 2, and it seeks to provide a minimum data rate, e.g., around 12 kbps. (Of course, keep in mind that a wireless network cannot provide an absolute guarantee. "Guarantee" here means that if the network agrees to grant service with QCI = 1 for a user, it will try its best to honor the granted GBR (Guaranteed Bit Rate). In the worst-case, the call may drop due to a hostile radio environment.[11])

Now, let's contrast QCI = 1 with QCI = 8. An application such as email might fall into QCI = 8. Since VoIP has more stringent delay requirements than email (e.g., 100 ms for VoIP vs. 300 ms for email), its priority is higher than email's. Also, the target error rate for email is lower than that for VoIP because the integrity of email bits is much more critical than the integrity of VoIP bits. So our goal is to lose no more than one of one million IP packets for email.

A web browsing session typically uses QCI = 8 or 9, which corresponds to so-called best-effort service. However, according to the 3GPP recommendations, the Evolved Packet System bearer with QCI = 8 or 9 is a non-guaranteed bit rate (non-GBR) bearer and therefore has absolutely no guarantee of any minimum data rate. The maximum data rate for such Evolved Packet System bearer is operator-configurable. Commercial LTE networks determine the maximum data rate as a function of the mobile device category. Since there is no guarantee of any minimum data rate for the best-effort service, the most critical parameter (i.e., throughput or data rate) cannot be specified for the best-effort service. Furthermore, the priority of traffic associated with the best-effort bearer is the lowest among all types of Evolved Packet System bearers. Hence, when the eNodeB scheduler becomes busy serving higher-priority bearers, the average throughput can be expected to be impacted for the best-effort bearers.

In summary, the absence of the specification within wireless standards of even the minimum data rate for a best-effort service is a hurdle that cannot be overcome by the transparency rule.

In light of the practical issues described above in Sections 4.1.1 through 4.1.5, applying an enhanced transparency rule to mobile wireless networks is impractical, would stifle innovations, and (most importantly) would not benefit users at all. Even if some information about network management practices were to be disclosed to comply with the transparency rule, such information would most likely be too vague. Metrics for gauging network neutrality do not exist and if they did they would likely become obsolete quickly with the rapid development of technology and new applications. Enforceability of such a rule would be highly questionable and this rule would, in practice, reduce network performance.

[11] The packet delay is the one-way time between the device and the edge of the operator's network. QCI = 1 aims for a delay of less than 100 ms. (The lower the number for priority is, the higher the actual priority.) The packet loss rate of 10^{-2} = 0.01 or 1% means that an application with QCI = 1 can tolerate the loss of one of 100 packets.

4.2 No-Blocking Rule and Mobile Wireless Networks

The *no-blocking rule* specifies that mobile broadband service cannot prevent consumers from accessing lawful websites and cannot prevent users from using voice or video telephony applications that compete with the mobile broadband service provider's services, subject to reasonable network management. The NPRM further clarifies that mobile broadband service providers would not be violating the rule if they do not degrade a lawful service or content below the minimum level of service. The NPRM is seeking a definition for such minimum level of service and exploring the feasibility of using measurements such as speed, packet loss, and latency delay to quantify the minimum level of service.

Application of the no-blocking rule has several unique challenges in mobile wireless networks. To begin with, the definition of the minimum level of service is not feasible for mobile wireless networks. Throughput is the most important performance metric for data-centric mobile broadband systems, and, as explained in Section 4.1.5 on the best-effort service, LTE does not define any minimum data rate guarantee for such service. Note that non-IMS applications such as consumer-chosen voice and video applications do not travel through the IMS network and are typically placed onto the Evolved Packet System bearer with QCI = 8 or QCI = 9. Recall from Section 2 that signaling for IMS applications such as operator-aware VoIP (e.g., Voice over LTE or VoLTE) travel through the IMS network and that the IMS and Policy and Charging Control networks work with the Evolved Packet Core to provide target QoS, which would include guaranteed bit rate (GBR) for QCI = 1 Evolved Packet System bearer. Hence, any non-IMS user application such as voice and video cannot be expected to have the IMS application-like QoS.

The issues associated with expansion of the no-blocking requirements for mobile wireless networks are similar to those addressed in Sections 4.1.1-4.1.5, above, for implementation of an enhanced transparency requirement. When a single user or single application could overwhelm the limited resources provided to wireless providers, such a provider must be able to block this interfering use to ensure the quality of service expected for many other users. As has been discussed throughout this paper, unlike wireline networks, mobile wireless networks have scarce spectral resources (capacity) that are affected by interference, multipath, blockage, clutter and other conditions which require active management, including blocking of particular applications or users. Without the ability to manage blocking effectively, a wireless provider would be faced with situations where a single user or application could occupy all the radio resources associated with a particular eNodeB – leaving any other subscriber seeking access to that eNodeB without the ability to connect and receive the service expected. Therefore, the current no-blocking regulation continues to be the most appropriate technical path forward. Attempting to apply a broader no blocking rule—even with a safe harbor set of guidelines or other means to cabin off "reasonableness" – is extremely impractical, as discussed in more detail below.

In case of resource crunch, the eNodeB gives higher priority to Evolved Packet System bearers carrying IMS signaling and guaranteed bit rate traffic (e.g., VoIP traffic). Furthermore, according to the Quality of Service Class Indicator characteristics defined in [3GPP_TS23.203], Evolved Packet System bearers are set with a certain Allocation and Retention Priority (ARP), and, by design, best-effort bearers have a lower Allocation and Retention Priority compared to other higher-priority bearers. The best-effort Evolved Packet System bearers carrying email, web browsing, and consumer-installed non-IMS voice and

video applications could potentially be affected adversely as part of routine and legitimate network management.

Without the differentiation capabilities described above, the LTE network will simply not function reliably for some services, or will function in a very inefficient manner. Standards organizations such as 3GPP have spent years working out the details of these capabilities and how they will interoperate, and they should not be modified without thorough technical analysis.

The NPRM is seeking comment on the feasibility of using the following methods to define a *minimum level of service*: a best-effort standard, a minimum quantitative performance standard, and a reasonable person standard. As discussed in Section 4.1.5, a best-effort standard cannot be really quantified for mobile wireless networks. The minimum quantitative performance standard would also be impractical. Finally, the main problem with the reasonable person standard is that a typical end user cannot be expected to be knowledgeable about how mobile wireless networks operate and different people would have different expectations from their networks. The absence of reliable and quantifiable estimates of "reasonableness" makes the reasonable person standard highly subjective and non-enforceable. Wireless providers, based on network management requirements developed within industry standards, should have the right to block any use or application on their wireless network if such use would preclude other subscribers from accessing service. Attempting to limit wireless providers' ability to block (such as attempting to define "reasonableness") would not allow the scarce spectral resources available to wireless providers to be used in the most effective and efficient manner.

4.3 Unreasonable Discrimination Rule and Anti-Discrimination/Commercial Reasonableness Rule, and Mobile Wireless Networks

The FCC has stated that the newly-proposed anti-discrimination/commercial reasonableness rule, just like the original 2010 rule, is not intended to be applicable to mobile broadband service. Our view concurs with the FCC view that different treatment for mobile broadband should be continued because differentiation among users and user services is required to provide a satisfactory quality of service to consumers.

As discussed in detail in Section 3.1 above, wireless networks are characterized by: (1) scarce radio resources; (2) radio resource sharing; (3) dynamic channel conditions and varying performance; (4) varying resource consumption; (5) ever-evolving networks and (6) the need to integrate differing devices and infrastructure. Because of these factors, user differentiation due to the dynamic nature of the radio environment is fundamental to the operation of any good scheduling algorithm design for a wireless network. A good scheduling algorithm maximizes network performance while providing good user-perceived experience, not necessarily by treating all users or all applications identically. If the scheduler treats two users with two different channel conditions (e.g., one excellent channel and one poor/noisy channel) in the same manner, the overall network performance would certainly degrade and the average user experience would also deteriorate. Consider Figure 5, where two users are downloading an email with a huge attachment and their channel conditions are constantly changing. Good channel conditions can support a higher data rate, and poor channel conditions support a lower rate as illustrated in Scenario 1 and Scenario 2.

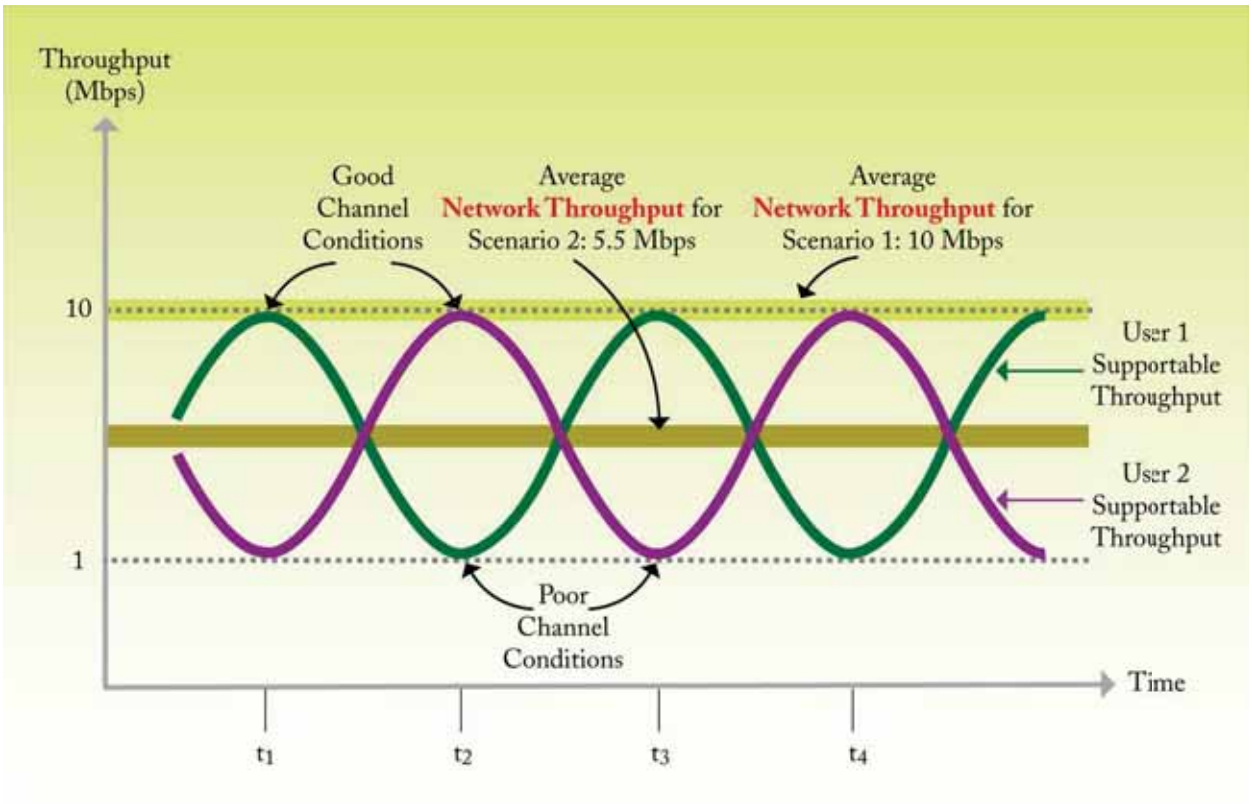

Figure 5. Necessity of User Discrimination Due to Dynamic Radio Environment

Figure 5 shows the user supportable throughput when all the available resources are allocated to the users. In Scenario 1, a high-performance scheduler allocates all the available resources to a user with the best channel conditions and transmits packets to such user. Observe that at time t_1, User 1 has the best channel conditions and can support 10 Mbps if allocated all resources. The scheduler dedicates the entire 100% of network resources to User 1 and sends a packet to User 1 at 10 Mbps at time t_1. At time t_2, User 2 has better channel conditions, and the scheduler allocates all network resources to User 2 and sends a packet to User 2 at 10 Mbps. The average network throughput is 10 Mbps as the network is always sending the packets at 10 Mbps. Sometimes the network sends packets to User 1, while other times, the network sends packets to User 2. The average user throughput that User 1 experiences is 50% of 10 Mbps = 5 Mbps, and the average throughput User 2 experiences is also 50% of 10 Mbps = 5 Mbps because these users are scheduled 50% of the time.

In Scenario 2, an equal-opportunity scheduler equally distributes the network resources at all times. At time t_1, the network allocates 50% of resources to User 1, leading to User 1 throughput of (50% of 10 Mbps = 5 Mbps). Note that User 1 throughput is 5 Mbps and not 10 Mbps because User 1 is allocated just 50% (and not all 100%) of resources. Similarly, at time t_1, the network allocates 50% of resources to User 2, leading to User 2 throughput of (50% of 1 Mbps = 0.5 Mbps). The network throughput at t_1 is 5.5 Mbps (User 1 throughput + User 2 throughput = 5 Mbps + 0.5 Mbps = 5.5 Mbps). Now, consider time t_2, where the allocation of 50% of resources to User 1 results in User 1 throughput of (50% of 1 Mbps = 0.5 Mbps) and the allocation of remaining 50% of resources to User 2 results in User 2 throughput of (50% of 10 Mbps = 5 Mbps). Again, note that the users experience only 50% of the throughput values shown

in Figure 5 because the throughput values correspond to a hypothetical case where all of the network resources are allocated to a single user. The network throughout at t_2 is (User 1 throughput + User 2 throughput = 0.5 Mbps + 5 Mbps = 5.5 Mbps). The average network throughput is then 5.5Mbps. Let's calculate average user throughput. User 1 experiences 5 Mbps 50% of the time and 0.5 Mbps remaining 50% of the time, leading to the average user throughput of 2.75 Mbps (0.5*5 Mbps + 0.5* 0.5 Mbps = 2.75 Mbps). Similarly, the average user throughput for User 2 is also 2.75 Mbps. In other words, since the network equally distributes resources between the two users, the network throughput of 5.5 Mbps is equally divided between the two users as (5.5 Mbps/2 = 2.75 Mbps).

In our simple example, the network throughput is reduced by almost 50% (i.e., from 10 Mbps to 5.5 Mbps) in Scenario 2 compared to Scenario 1. Just imagine what would happen to the business models of service operators if the cost of supporting their customers doubles overnight? While the scheduler has optimized network performance in Scenario 1, User 1's throughput and User 2's throughput are also better in Scenario 1 compared to Scenario 2 (e.g., 5 Mbps in Scenario 1 compared to 2.75 Mbps in Scenario 2). Better network performance enables the service operator to cost-effectively provide services to many users simultaneously. Subscription plans for users can then be relatively inexpensive, promoting growth of cellular subscribers and services. The comparison of network performance in Scenarios 1 and 2 shows that differentiation is best for the aggregate network and for all users. Treating all users the same all the time appears more fair at first, but adapting to the radio channel by having the scheduler weight the service schedule against predicted data delivery performance results in better performance for everyone, even though at any moment, not all users are treated the same.

Combined service and user differentiation is also quite important. Assume that User 1 has an ongoing email application and has in the past been promised a maximum data rate of 10 Mbps, and assume further that all the network resources are being consumed by such a user. Suddenly, ten users start making voice calls. The network simply lacks the resources to simultaneously support ten voice users and an email user with a 10 Mbps data rate. If the network's resource management algorithms downgrade the email data rate to perhaps 9 Mbps, then the network can accommodate both the email user and all ten voice calls. If the network fails to differentiate between the voice users and the email user, all ten voice calls would be blocked. In summary, user and service differentiation is essential to aggregate service fairness for the average consumer. Here again, the scheduler is not treating every application identically ('fairly'), but the net aggregate result benefits more users. What set of metrics would represent this fairness? These considerations evolve as the applications mix changes.

Differentiation based on resource consumption is also inherent in a wireless network and facilitates network efficiency and fairness. The network management algorithms must differentiate between users based on the amounts of network resources each user is consuming. For example, current mobile wireless networks commonly limit the amount of resources a single user can consume. If one user consumes an excessive amount of network resources due to a hostile radio environment and/or such user is using bandwidth-intensive data applications, that user may dominate the network so much that no other user can get any service in the absence of pro-active network management.

There are several situations where it is legitimate and beneficial for wireless network operators to differentiate traffic. For example, Wireless technologies are increasingly being used for machine-to-machine services and public welfare systems. It is critical that these systems – such as, for example, wireless monitoring of bridges – be fully functional at all times, and this may require prioritization. In

addition, public safety personnel clearly should have higher priority than regular users. More "ordinary" services, such as voice call, email, and streaming video, all require different quality of service levels, and wireless network operators should be allowed the flexibility to prioritize these diverse services in a manner that ensures that an end user experiences the quality of service necessary for these services to function. User and service differentiation is also essential to service fairness – one user should not be permitted to monopolize network resources at the expense of others.

In summary, the dynamic nature of mobile wireless networks requires a reasonable, necessary, and *dynamic*differentiation among users by the network management to ensure an acceptable aggregate quality of service for all wireless subscribers. To subject mobile broadband providers to claims that such non-uniform network management techniques are 'unfair' and violate commercial practices, particularly when combined with the prospect of regulatory rebuke, would significantly impair the ability of providers to experiment with new and innovative network management tools designed to improve consumers' experiences. Any rule that would prohibit discrimination on mobile wireless networks would be impractical and would actually work against the FCC's goals of promoting innovation and benefiting consumers.

5. Mobile Wireless Broadband Internet Access: An Integrated Information Service

Mobile broadband service is an integrated service that enables the wireless subscriber to access a variety of services in a wireless fashion. The subscriber's device communicates with the mobile broadband service provider's network via complex interactions. The nodes of the entire wireless network infrastructure work together to present a single unified view of the network to the subscriber's device and to provide service-specific QoS for a user's services according to the 3GPP LTE framework. All the network components need to do specific processing, which often needs to be customized for a given service, to provide seamless and satisfactory experience of a variety of services for the user. All the complexities associated with subscriber's experience of wireless services are handled by the subscriber's device and the broadband service provider's network without the active involvement of the subscriber.

When the FCC classified wireless broadband Internet access service as an "information service," it did so based on the correct finding that this service "offers a single, integrated service to end users, Internet access, that inextricably combines the transmission of data with computer processing, information provision, and computer interactivity, for the purpose of enabling end users to run a variety of applications."[12] This statement, which was made by the FCC in 2007, has only become more emblematic of the wireless ecosystem. As technologies and networks have evolved, subscribers are increasingly using advanced networks for multiple simultaneous data services, such as email, web browsing, and various other applications. Extensive and complex processing in the mobile broadband network allows customers to seamlessly navigate among multiple mobile broadband applications and services at the same time, enjoying a good experience of various applications.

The mobile broadband network consists of numerous network nodes that interact among themselves in different and complex ways and that do custom processing depending upon the type of service or application. Such interactions and custom processing enable the wireless subscriber to obtain an

[12] *Appropriate Regulatory Treatment for Broadband Access to the Internet Over Wireless Networks,* Declaratory Ruling, 22 FCC Rcd 5901, 5911 ¶ 26 (2007).

integrated information service that integrates different types of information to provide a unified service experience (user experience) and that meets specific requirements of applications (e.g., guaranteed data rate or very low packet error rate or very low latency). *Indeed, the mobile broadband service is an integrated information service that requires (i) tight coupling between the mobile device and the network,*[13] *(ii) numerous complex interactions*[14] *between the mobile device and the network and among the network components, and (iii) service-specific custom processing at different network nodes.*[15] Let's take a closer look at these three areas.

Tight coupling between the mobile device and the network is essential in providing seamless and satisfactory services to the subscriber. For example, each service requires a certain quality of service, and, the network properly configures the mobile device and the network nodes so that the user has satisfactory experience. According to the 3GPP LTE standard, the overall packet error rate cannot be greater than 0.0001% for services such as email and web browsing (see Table 6.1.7 in 3GPP TS 23.203). However, the raw packet error rate on the LTE air interface is 10%.[16] Hence, the network configures a suitable number of packet retransmissions to reduce the effective packet error rate from 10% to 0.0001%. Furthermore, the mobile device provides feedback on the prevailing downlink radio channel conditions so that the network can use suitable transmission parameters (e.g., the modulation scheme, the amount of redundancy, the type of multiple antenna technique, and the number of Physical Resource Blocks) to provide a satisfactory downlink data rate and hence a satisfactory user experience. Similarly, the mobile device informs the network about the amount and type of data it has in its uplink buffers and the available transmit power. The network allocates a suitable amount of uplink radio resources based on such information and the subscriber can send the data traffic (e.g., email) within acceptable delay limits (e.g., less than 0.3 second). This tight coupling enables end users to receive email, for example, at a data rate that would be expected with very limited errors. Without this network management, the quality of service would deteriorate and be unacceptable to subscribers.

Complex interactions between the mobile device and the network and among the network components take place before the subscriber can obtain even basic wireless services. Mobile devices typically do not have pre-assigned fixed IP addresses, and, the devices cannot obtain any IP-based services such as email and web browsing without IP addresses. Hence, the network *must* allocate an IP address to the mobile device. To provide security over the wireless interface, the network and the device first perform mutual authentication and then locally generate security keys. For example, LTE can secure the wireless interface by encrypting user traffic. The network also sets up several logical connections called Evolved Packet System bearers that help carry user traffic such as email and streaming video. The network

---

[13] Such tight coupling is exemplified by packet retransmissions occurring between the user device and the radio network to provide essentially error-free information to the applications such as e-mail and web browsing.

[14] An example of such interaction is the invocation of a Domain Name System (DNS) server by the user device so that the name of a web site (e.g., www.cnn.com) can be translated into an IP address of the server that is in charge of the web site.

[15] An example of such custom processing include fast packet forwarding of delay-sensitive traffic by an IP router and delayed packet forwarding of delay-tolerant traffic by an IP router. The operator's network utilizes multiple IP routers within the network (e.g., between the Serving Gateway and the Packet Data Network Gateway in the LTE network).

[16] The LTE air interface uses the instantaneous target block error rate (BLER) of 10% to improve efficiency of precious radio resources. A suitable combination of the modulation scheme and Turbo coding is used to meet such target BLER for a given radio channel condition.

nodes interact among themselves and the network interacts with the mobile device so that the bearers can be set up. Selected information about the mobile device is stored at different nodes so that packets can reach the correct user via the bearers. Without all these integrated actions, the user would not be able to obtain the Internet services expected (i.e., would not be able to access the desired web site). The wireless provider must manage these complex interactions to provide the seamless experience expected by consumers.

Service-specific custom processing is carried out at different network nodes. Depending upon the policies of the service provider, different types of IP addresses could be allocated to the mobile device for different packet data networks. For example, for the packet data network of the Internet, an IPv4 address could be allocated to the mobile device because of prevalence of IPv4 addresses. In contrast, for IMS-based applications, an IPv6 address could be allocated to the mobile device to benefit from the abundance of IPv6 addresses. Quality of Service (QoS) in an IP network can be provided by an Integrated Services or Differentiated Services framework. The network node provides different QoS to different services to meet the service requirements and user expectations. When a Differentiated Services framework is used, Differentiated Services Code Point (DSCP) is used to mark each IP packet. IP routers use correct packet forwarding treatment to an incoming IP packet. For example, assume that two IP packets arrive at a Packet Data Network Gateway: a delay-sensitive IP packet carrying a streaming video and a delay-tolerant IP packet carrying email. The delay-sensitive IP packet carrying a streaming video can be marked with the DSCP value of 30 and the delay-tolerant IP packet can be marked with the DSCP value of 0. In case of heavy traffic, the IP routers between the Packet Data Network Gateway and the Serving Gateway would quickly forward the delay-sensitive packets (i.e., video streaming packets) and would delay the forwarding of the delay-tolerant packets (i.e., email packets). The IMS and the PCC network nodes also work with one another such that the bearers can help meet different QoS requirements for different services. This in turn allows the wireless provider to ensure that subscribers that are not affected adversely by latency to be delayed, while those applications that are latency sensitive are not delayed. For example, video streaming would not be delayed so that playback is acceptable for a subscriber, while email packets could be marginally delayed but consumers would not be affected by this delay. This network management allows the provider to manage the scarce spectrum resources in an efficient, effective manner, without degrading the subscriber experience.

Close cooperation between the mobile device and the network is needed for cohesive and seamless integrated service experience for the wireless subscriber.

Examples of Integrated Wireless Broadband Services

The tight integration needed to provide wireless broadband services is demonstrated below by how a consumer obtains two services, web browsing and video conferencing. For both of these cases, the mobile device and the broadband network must work together to provide a seamless and integrated service experience for the consumer. Before any services are rendered to the consumer, the mobile device synchronizes with the radio network and performs an attach operation with the network. As part of the attach operation, mutual authentication occurs, and security between the mobile device and the network is established. Furthermore, two default bearers are established, one for the Internet packet data network and one for the IMS Packet Data Network. The Packet Data Network Gateway allocates to the mobile device an IP address for the Internet Packet Data Network and a separate IP address for the IMS Packet Data Network. Additionally, IP addresses of the DNS server and the IMS server are conveyed

to the mobile device. When the consumer initiates video conferencing, additional bearers[17] are established to carry voice and video through numerous interactions among the mobile device, the radio network, the core network, the IMS network and the PCC network. Now that all the groundwork for data traffic has been completed, the IP packets for the email and video conferencing packets start flowing through the network and the consumer reaps the benefits of all the hard work that the mobile device and network nodes have been doing.

Web Browsing on Mobile Broadband Networks

Let's summarize the *major steps involved when the consumer is browsing the web* as illustrated in Figure 6.[18]

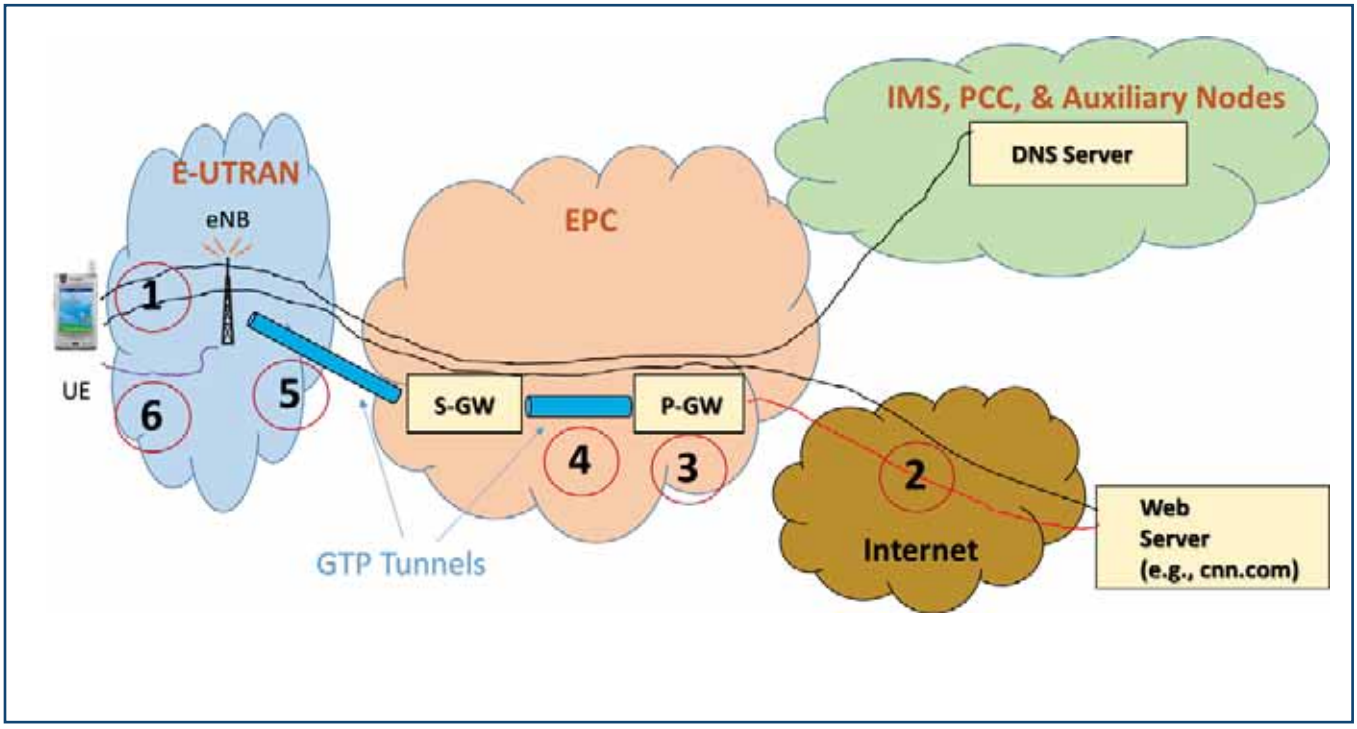

Figure 6. Major Communication Steps for Web Browsing

When the consumer selects www.cnn.com as the web site for browsing in Step 1, the mobile device communicates with the DNS server to find the IP address of the web server in charge of www.cnn.com and then communicates with the web server to set up an end-to-end connection with the web server. In Step 2, the web page from the web server passes through the IP routers of the Internet and arrives at the Packet Data Network Gateway. The Packet Data Network performs translation between the public IP address and the private IP address (if needed) in Step 3. In Step 4, the Packet Data Network

[17] LTE controls QoS at the levels of the EPS bearers. Hence, two bearers are required for two services that need two different QoS levels.

[18] The simplified description here represents one possible approach for providing web browsing and video conferencing services. The LTE standard is quite flexible and operators can choose a variation of the approach described here to offer services to the user. In the interest of simplicity, only selected nodes and connections are shown in the figure.

determines the correct bearer for the incoming IP packet containing the web page and places the IP packet inside a tunnel (representing part of the bearer) toward the Serving Gateway using a protocol called GTP.[19] The Serving Gateway extracts the IP packet and places it inside another tunnel toward the eNodeB using GTP protocol in Step 5. The eNodeB uses several protocols of the air interface protocol stack to format the original IP packet for air interface delivery (including encryption for security) in Step 6 and then transmits the web page packet over the air to the mobile device. In Step 7, the mobile device also uses the protocols of the air interface protocol stack to extract the original IP packet from the air interface (including decryption) and then presents the web page to the consumer. *As is evident from these steps, the mobile broadband network nodes and the mobile device work closely together to present web browsing to the consumer as an integrated information service*.

Video Conferencing on Mobile Broadband Networks

Figure 7 summarizes the main steps involved when the consumer is participating in video conferencing [Radisys_October2012] [Ericsson_ Feb2012].

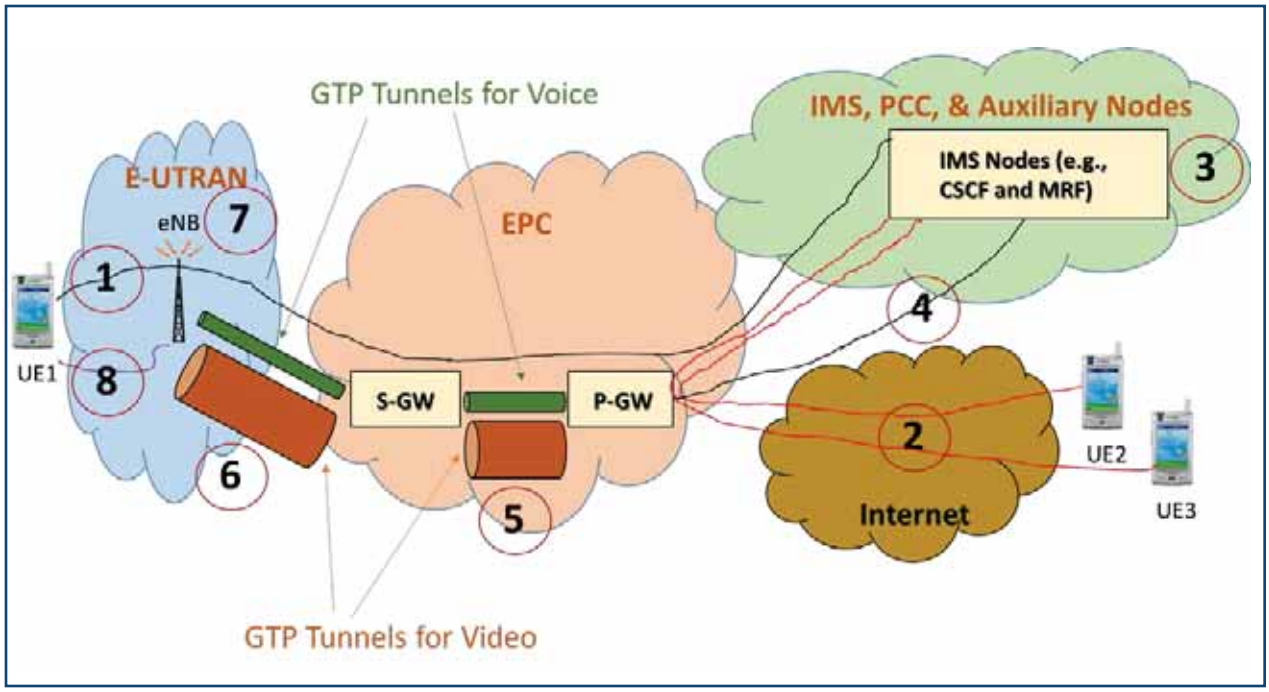

Figure 7. Major Communication Steps for Video Conferencing

[19] GTP stands for GPRS Tunneling Protocol, where GPRS is General Packet Radio Service. GTP enables LTE to support IP mobility, where the user can move from one geographic location to another while maintaining the same IP address. IP mobility is one of key elements of a mobile broadband service.

When the consumer chooses a video conferencing application in Step 1, the mobile device 1 communicates with the IMS network to create communication paths among the participating devices. In Step 2, the voice and video media from other devices such as mobile device 2 and mobile device 3 arrive at the Media Resource Function in the IMS network. The Media Resource Function properly mixes the voice and video media streams (and performs media codec conversion if needed) in Step 3. In Step 4, the Media Resource Function sends the IP packets containing voice and video to the Packet Data Network Gateway. The Packet Data Network Gateway in Step 5 determines the correct bearers for the incoming IP packets, places the voice packets inside one tunnel toward the Serving Gateway using GTP, and places the video packets inside another tunnel toward the Serving Gateway using GTP. The Serving Gateway extracts the IP packets and places them inside other tunnels toward the eNodeB using GTP in Step 6. The eNodeB uses several protocols of the air interface protocol stack to format the original voice and video IP packets for air interface delivery in Step 7 and then sends these IP packets over the air to the mobile device in Step 8. In Step 9, the mobile device 1 uses the help of the air interface protocols to extract the original voice and video packets from the air interface and then plays voice and video to the consumer in the video conferencing application. *As is evident from these main steps, the mobile broadband network nodes and the mobile device work very closely together to offer video conferencing to the consumer as an integrated information service*.

See Appendix I for a more detailed discussion of how the mobile broadband network provides an integrated information service to subscribers.

6. Recommendations

Mobile wireless broadband networks face unique challenges such as the scarcity of spectrum and other radio resources, dynamic radio environment, varied and changing technologies for devices, base stations, and networks, differentiated and evolving services and applications, and exponentially-rising data traffic. Based on the in-depth analysis of the modern mobile wireless networks, we respectfully recommend the following to the FCC.

- ✓ Mobile wireless broadband networks must be treated differently from other communications networks such as fixed wireless networks and wireline networks.
- ✓ Aim for minimal regulations for mobile wireless networks to promote innovations and thereby facilitate achieving the ultimate goals of superior network and spectral efficiency and excellent user experience.
- ✓ Give maximum flexibility to the design and optimization of complex and distributed wireless network management so that the networks operate with maximum possible efficiency under the constraint of limited spectrum.
- ✓ Refrain from establishing any minimum performance standards, because these standards are simply impractical to define or enforce in mobile wireless networks.
- ✓ Ensure that proprietary and competitive management processes are respected and encouraged to motivate continuing innovation and differentiation.

References

[NPRM_2014] FCC, Notice of Proposed Rulemaking, “Protecting and Promoting the Open Internet,” GN Docket No. 14-28, May 15, 2014.

[Neel_MobileDataTraffic] James Neel, "Emerging Wireless Standards," Wireless Summer School, http://www.crtwireless.com/EWS_10.html, June 2010.

[Deloitte_SpectrumShortage] Deloitte, “Airwave overload?,” http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/TMT_us_tmt/us_tmt_Spectrum_Thought_Leadership_September_092512.pdf, September 2012.

[Ericsson_February2012] Ericsson, “Voice and Video Calling over LTE,” February 2012, http://www.ericsson.com/res/docs/whitepapers/WP-Voice-Video-Calling-LTE.pdf .

[Radisys_October2012] Adnan Saleem, Radisys, “Leveraging IMS for VoLTE and RCS Services in LTE Networks,” ETSI Workshop- RCS, VoLTE, and Beyond, Kranj, Slovenia, October 2012, http://www.etsi.org/plugtests/RCS-VOLTE/pres/Radisys-ETSI-MSF_Workshop_RCS_VoLTE_MRF.pdf .

Appendix I

A Closer Look at the Mobile Broadband Internet Access Service as an Integrated Information Service

Let's dive into the details of how the mobile broadband network provides an integrated information service to the subscriber.[20] Consider Figure 8, where the wireless subscriber is using an LTE network for three simultaneous services- web browsing, email, and video conferencing.[21]

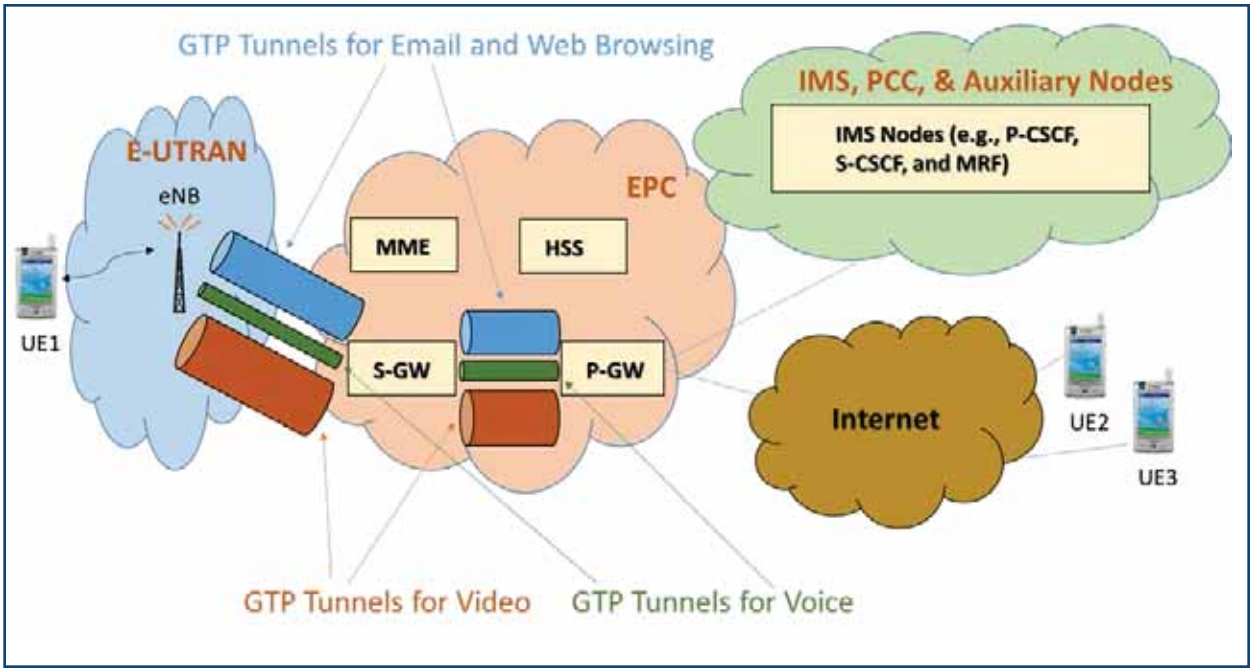

Figure 8. Integrated Information Service Offered by the Mobile Broadband Network

Different network nodes process different aspects of signaling and/or traffic. The following steps are executed to provide the integrated information service to the wireless subscriber: (I) initial attach and default EPS bearer setup toward the IMS network, (II) default EPS bearer setup toward the Internet, (III) dedicated EPS bearer setup toward the IMS network, and (IV) data transfer. Let's take a closer look at these steps below. *These steps show that complex interactions and node-specific custom processing are instrumental in providing an integrated information service experience to the wireless subscriber*.

Initial Attach and Setting up of the Default EPS Bearer Setup toward the IMS Network

As summarized in Section 2.2, the UE carries out the initial attach procedure upon power-up. The UE synchronizes with the eNB and establishes an RRC (Radio Resource Control) signaling connection with a cell. This RRC connection helps the UE and the eNB exchange signaling messages such as the messages

[20] The description given here is one possible implementation. The LTE standard is quite flexible and the operator can choose a variation of the approach

[21] This figure is based on the information given in [Ericsson_February2012] and [Radisys_October2012].

related to handover and EPS bearer setup. The UE and the MME perform mutual authentication. The HSS helps the MME in the authentication process. The signaling connection between the UE and the MME is a NAS (Non-Access Stratum) connection, which can be used for messages such as EPS bearer setup messages. Security is activated for the RRC connection and the NAS connection. The responsibility of securing the air interface between the UE and the eNB lies with the UE and the eNB. The responsibility of securing the NAS signaling connection between the UE and the MME lies with the UE and the MME. The service provider determines which packet data network should be used as the default packet data network. In Figure 6, the Internet and the IMS network are potential candidates as the default network. In our scenario, the network establishes a default EPS bearer with QCI = 5 (which is recommended by the 3GPP [3GPP_TS23.203]) that has the highest priority among the QCIs. During the default EPS bearer setup, the P-GW allocates an IP address to the UE, and, this address is conveyed to the UE by the MME. The MME also conveys to the UE IP addresses of the P-CSCF (Proxy- Call Session Control Function) and the Domain Name System (DNS) server. The P-CSCF is the first point of contact of the UE with the IMS network and performs various functions such as compression of SIP signaling messages and interactions with the PCC to provide end-to-end QoS for IMS-based services. The DNS server is used by the UE to resolve the website names to the IP addresses so that the UE can exchange IP packets with websites. The default EPS bearer toward the IMS network helps the UE and the network exchange SIP signaling messages.

Setting up of the Default EPS Bearer Setup toward the Internet

After the default EPS bearer with the default packet data network (e.g., the IMS network in our scenario) is set up, a different default EPS bearer with QCI-8 or 9 toward the Internet is established. The P-GW allocates an IP address to the UE for the Internet access. This default EPS bearer helps carry traffic corresponding to applications such as email and web browsing.

Setting up of the Dedicated EPS Bearers toward the IMS Network

As discussed earlier in Section 3, LTE controls the QoS at the level of EPS bearer. Although a default EPS bearer toward the IMS network has already been established, the QoS of this EPS bearer is inadequate to carry voice and video traffic. The QCI = 5 EPS bearer is a non-GBR bearer, while the voice and video need data rate guarantees for satisfactory user experience. When a user initiates video conferencing, SIP signaling messages are exchanged between the UE and the IMS network. The P-CSCF observes such signaling messages and conveys the information about the QoS requirements of the call to the Policy and Charging Rules Function (PCRF). The PCRF consults Subscription Profile Repository to check if the QoS requested by the video conferencing application can be accepted. The PCRF translates the generic QoS description extracted from SIP signaling messages into the LTE-specific QoS parameters (e.g., a numerical values for QCI and determination of the GBR) and conveys the LTE-specific QoS to the Policy and Charging Function (PCEF). According to 3GPP, P-GW acts as the PCEF. The P-GW initiates the setup of (i) the dedicated EPS bearer that can carry the voice traffic with suitable QoS and (ii) the dedicated EPS bearer that can carry the video traffic with suitable QoS. The S-GW conveys the dedicated EPS bearer requests to the MME. The MME works with the eNB and the S-GW to set up network resources for the new dedicated EPS bearers. The eNB accepts the dedicated EPS bearer requests if it has adequate radio resources to support the QoS required for the video conferencing service.

The IMS network, and in particular, the Media Resource Function (MRF) plays an important role in supporting the video conferencing service. The MRF consists of the signaling entity called Media Resource Function Controller (MRFC) and the user traffic entity called Media Resource Function Processor (MRFP). The MRFC works with other IMS entities and facilitates the establishment of communication paths among the devices that are participants of video conferencing. The MRFP is responsible for mixing voice streams and video streams so that the mobile device can receive audio and video from all other participants.

Data Transfer

Consider the web browsing service. When the subscriber enters the website name in the browser (e.g., Internet Explorer or Chrome), the mobile device contacts the DNS server to receive the IP address of the website. The mobile device and the web server can now exchange IP traffic such as web pages and acknowledgements to the received IP packets. The IP packets containing the web page from the web server pass through the Internet routers and arrive at the P-GW that gave the UE the IP address associated with the Internet. In case a private IP address had been assigned to the UE, the P-GW translates the public IP address into a private IP address for the journey of the IP packet within the operator's LTE network. The P-GW places the IP packet on the GTP tunnel (associated with the default EPS bearer for the Internet) toward the S-GW. The S-GW removes the IP packet from the P-GW side of the GTP tunnel and places the IP packet on the GTP tunnel toward the eNodeB. The eNodeB extracts the IP packet and passes it through these protocols of the air interface protocol stack so that IP packet is in the format that can survive the hostile and dynamic radio environment- Packet Data Convergence Protocol,[22] Radio Link Control,[23] Medium Access Control,[24] and Physical Layer.[25] The eNodeB then transmits the formatted IP packet over the air. The mobile device acquires the formatted IP packet from the air interface and recovers the original IP packet by using the air interface protocols. Finally, the IP packet is made available to the web browser that displays the actual content to the subscriber.

Let's focus on the video streaming service now. After the video conferencing service has been initiated and the IMS network has helped establish communication paths among the participants, the voice and video traffic can start flowing. Voice and video streams from the participants of the video conferencing arrive at the MRF. The MRF mixes the audio and video streams of the participants and sends the IP packets carrying voice and video to the P-GW. The P-GW has two dedicated bearers with the UE in support of video conferencing, one for voice traffic and one for video traffic. The IP packets containing voice are placed onto the tunnel associated with the voice traffic and the IP packets containing the video are placed onto the tunnel associated with the video traffic. Once the S-GW retrieves the IP packets from the P-GW side of the tunnels, it removes the IP packets from the P-GW side of the GTP tunnels and places the IP packets on the GTP tunnels toward the eNodeB. The eNodeB extracts the IP packets and passes them through the layers of the air interface protocol stack mentioned above. The eNodeB then transmits the formatted IP packets over the air. The mobile device retrieves the formatted IP packets from the air interface and recovers the original IP packets carrying voice and video by using the air

[22] Example functions of Packet Data Convergence Protocol are encryption and header compression.
[23] Example functions of Radio Link Control are in-sequence delivery of packets and retransmissions of erroneous packets.
[24] Example functions of Medium Access Control are scheduling and management of radio resources and control of physical layer retransmissions.
[25] Physical Layer takes care of functions such as modulation, coding, power control, and multiple access.

interface protocols. Finally, the IP packets are sent to the video conferencing application that plays the voice and video content for the subscriber.

The transport network that carries the traffic between eNB and the S-GW and between S-GW and the P-GW also needs to provide different QoS treatments to best-effort traffic such as email and web browsing and delay-sensitive traffic such as voice and video packets. The mechanism such as Differentiated Services (DiffServ) is widely used for QoS in an IP network. DiffServ involves marking of the IP packet by a code called DiffServ Code Point (DSCP). Different values of DSCP are defined for different services so that the IP router that forwards the IP packet containing a given service (e.g., email, web page, voice, or video) toward the destination gives a suitable priority to the incoming IP packet. For example, DSCP=0 means that the service is a delay-tolerant and the IP router could let this IP packet wait in the buffer for some time when it is busy with IP packets carrying other higher priority services. In contrast, DSCP=30 means that the IP packet is carrying a video streaming packet and hence warrants faster packet forwarding from the IP router.

As discussed in the paragraphs above, numerous and complex interactions among the mobile device, the radio network, the core network, the IMS network, and the PCC network are required so that the subscriber can access a variety of Internet access services. *Multiple services are offered to the consumer as a single integrated information service with the mobile device and the network working closely together and carrying out complex, intense, and custom processing*.

CTIA
The Wireless Association®

CTIA Perspective on Preserving an Open Internet

Net Neutrality and Mobile Broadband Networks

Key Points

- Competition in the U.S. mobile broadband market is delivering an open mobile Internet, providing world-leading mobile broadband networks, and creating an innovative mobile ecosystem that benefits U.S. consumers and businesses. Internet openness is essential for any carrier seeking to win and keep customers.

- Technical, operational, economic, developmental and other factors continue to make mobile wireless services different and warrant a mobile-specific regulatory approach.

- Invitations to apply backward-looking Title II public utility regulation are barred by statute, would undercut investment, stifle innovation, and give rise to great uncertainty, all of which would harm consumers.

The U.S. Mobile Wireless Ecosystem Has Thrived Since 2010

- **LTE Deployment**: The U.S. is the global leader in 4G deployment: with just 5 percent of the world's wireless connections, we have almost half of all LTE connections worldwide.
- **Investment**: Wireless network providers have invested over $90 billion in advanced wireless networks since 2010.
- **Apps Economy**: Analysts expect mobile app downloads worldwide to exceed 139 billion this year.

Mobile Trends in the United States, 2010–2013

	2010	2013	Increase
Apps	800 thousand	3.6 million (5.5 million as of May)	+347%
Cumulative Capex	$310 billion	$398 billion	+29%
LTE Subscribers	215,000	95.2 million	+44,168%
Data Traffic	388 billion MB	3.23 trillion MB	+732%
Video Traffic	24 PB a month	200 PB a month	+733%
Minutes of Use	2.24 trillion	2.62 trillion	+17%
Cell Sites	253,000	304,000	+20%
Smartphones	78 million	175 million	+124%
Tablets *(CMRS only, not Wi-Fi)*	13.5 million	25.2 million	+86%
Total Handset Models Available	630	790	+25%

CTIA, Press Release, CTIA's Annual Survey Says US Wireless Providers Handled 3.2 Trillion Megabytes of Data Traffic in 2013 for a 120 Percent Increase Over 2012 (June 17, 2014), available at http://www.ctia.org/resource-library/press-releases/archive/ctia-annual-survey-2013; Cisco, VNI Mobile Forecast Highlights 2013–2018, http://www.cisco.com/assets/sol/sp/vni/forecast_highlights_mobile/index.html#~Country (last visited July 13, 2014) (accessed by selecting United States, Mobile Applications); Informa Telecoms & Media Group, WCIS+ Database (last visited September 15, 2014).

Mobile Wireless Networks and Their Unique Technical Characteristics

- Scarcity of spectrum
- Dynamic radio channel conditions
- Need to share radio resources with varying quality of service (QoS) requirements
- Mobility
- Slower process of network capacity and growth
- Coexistence of devices and network technologies with widely different capabilities

Characteristics of Wireless Networks

Scarcity of Spectrum and Exponential Growth of Wireless Data → Network management needs maximum flexibility to maximize efficiency so as to serve all users

Change is Constant…multiple generations, revisions of a generation, & upgrades in a revision → Impractical to track network management mechanisms in a timely fashion for regulatory needs

Network Characteristics & Policy Implications

Complex & Dynamic Radio Channel, User Mobility, and Quality of Service → Requires discrimination and fast adaptation by network management

Dynamic Radio Channel, Ever-Changing Networks, Variety of Apps, & Shared Radio Resources → Impractical to quantify any truly verifiable performance guarantees

Inherent Differences between Wireline and Wireless Networks

Characteristic	Wireline	Wireless
Communications Channel	Relatively clean with signal regeneration	Impaired with noise, interference, multipath, and blockage
Bandwidth	No spectrum limitations	Severe Spectrum limitations
Mobility	None	Constant, complex, often unpredictable, and often consuming extensive resources
Security	A lesser concern due to the physical path between the provider and the user (buried or on aerial infrastructure).	A greater challenge due to the possibility of tracking a user and variety of interfaces
Response to Capacity Demand	Capacity increases may be feasible, although soaring demand and increasing congestion issues may call for additional pricing, bandwidth limitations, and prioritization mechanisms	Primarily managed dynamically through prioritization, scheduling, and power allocation.
Network Complexity	Relatively simple	Extremely complex
Network Stability, Deployment, and Maintenance	Comparative stable platform and systems, although high growth in demand and new applications are issues	Extremely dynamic platforms and systems; Deployment and maintenance require constantly dealing with real estate acquisition and zoning issues; Planning and maintenance are more difficult, and continuous maintenance and frequent resetting of network parameters required; Infrastructure changes to address localized capacity issues have ripple effects through adjacent cells.
Quality of Service	Easier to implement due to availability of higher capacity and predictability of resource requirements	Quite difficult to implement due to variable capacity, unpredictability of resource requirements, and existence of proprietary mechanisms; Industry moving towards IMS and PCC.

Differentiation: A MUST for the Wireless Network

User Differentiation: Channel Conditions

- Good channel: higher data speeds
- Poor channel: lower data speeds
- Network management needs to differentiate among users experiencing different channel conditions to increase aggregate speeds for ALL users

User Differentiation: Resource Consumption

- From excellent to poor channel conditions, resource consumption can increase by a factor of 36 for a user, starving other users
- Network management needs to differentiate among users to satisfy needs of all users

Differentiation

Cost of "No Differentiation"

More blocked calls and reduced data speeds

Service Differentiation: Quality of Service

- Different services such as web browsing and voice calls need different Quality of Service
- Network management needs to differentiate among user services to optimize overall user experience of these services

Transparency

- The Transparency Rule
  - The FCC should retain, but not expand, the existing transparency rule.
  - The proposed expanded transparency requirement of network performance including throughput, congestion and peak load management would be meaningless and potentially confusing to a subscriber as these metrics are monitored and changed on a millisecond by millisecond basis.

Transparency Rule & Wireless Networks

Network Management

o Proprietary, product differentiators: Disclosure would reduce innovations, and degrade user experience

o Always-evolving to reflect new revisions & lessons learned: Highly impractical to track
o Critical to network performance and user experience

Parameters of Default or Best-Effort Service

o No guarantee of main metric (i.e., data speed) per 3GPP
oMultiple technologies & ever-evolving network and device capabilities: Any potential quantification outdated quickly

Challenges of Additional Disclosures

Congestion Specifics

o Dynamic resource allocation and ongoing troubleshooting and optimization of network make quantification of congestion impractical

Performance

o Dynamic radio environment, resource allocation operating every millisecond, and varying QoS requirements of applications: Any real-world quantification of performance and its verification for regulatory compliance impossible

Expanded Regulation Would Harm Consumers

- An expanded No Blocking Rule would inhibit wireless providers from managing their networks. Determining a "Minimum Level of Service" for the variety of communications and circumstances within the wireless network is infeasible.
- An Anti-Discrimination/Commercial Reasonableness Rule would inhibit mobile providers from differentiating network traffic and introducing differentiated offerings in the marketplace.
- Applying a one-size-fits-all regime would reduce consumer options.

A Reasonable Network Management Exception Cannot Remedy Overbroad Rules

- Network management is the rule not the exception in mobile broadband
- Expanded rules would stifle innovation and limit consumer choice

Competition Delivers for Wireless Consumers

"The American consumer has been the beneficiary: new pricing and new services that have been spurred by competition."

-- Prepared Remarks of FCC Chairman Tom Wheeler, 20
14 CTIA Show, Las Vegas, NV (Sept. 9, 2014)

"Competition today is driving enormous benefits in the direction of the American Consumer."

-- Remarks of Bill Baer, Assistant Attorney General, Antitrust Division, U.S. Department of Justice (Jan. 30, 2014)

- Mobile Broadband Speeds: Increased 8 times over since 2010
- Mobile Broadband Prices: Dropped 93% between 2008 and 2012
- Consumer Choice: 92% of Americans are served by 3 or more mobile broadband providers (as compared to 15% for wired providers)

CTIA
The Wireless Association®

Title II Would Jeopardize the Mobile Marketplace's Dynamism, Deter Investment and Innovation, and Harm Consumers

- Section 332 bars the FCC from reclassifying mobile broadband into a common carrier service.
- A Title II regime would subject mobile broadband to a host of inappropriate regulations, eviscerating the ability to innovate and suppressing investment.
- Forbearance associated with a "third way" approach would itself be extremely risky.
- No matter how the Commission effectuated reclassification, it would be subject to years of litigation, extending uncertainty and harming consumers.

Mobile Broadband Is A Single Integrated Information Service

- FCC made correct determination in 2007 and 2010
- Tight coupling between the device and the network
- Complex interactions between the device & the network (e.g., invocation of a Domain Name System (DNS) server by the user device),
- Service-specific custom processing at different network nodes (e.g., fast packet forwarding of delay-sensitive traffic by an IP router)