Ixia Enabling a Converged World™

Quality of Service (QoS) and Policy Management in Mobile Data Networks

Validating Service Quality to Ensure Subscriber Quality of Experience (QoE)

915-2731-01 Rev B July 2011

Contents

Introduction

Cellular systems by nature have finite resources. Radio spectrum and transport (backhaul) resources are limited, expensive, and shared between many users and services. Mobile broadband networks must support multiplay applications of voice, video, and data on a single IP-based infrastructure. These converged services each have unique traffic-handling and QoE requirements. Such issues cannot be economically solved by over-provisioning the network. A positive user experience must be obtained through efficient partitioning of the available wireless network resources.

The 3rd Generation Partnership Project (3GPP) – the Universal Mobile Telecommunications System (UMTS) and Long-term Evolution (LTE) standards body – has developed a comprehensive QoS, charging, and policy control framework to address this problem. The policy and charging control (PCC) is the heart of the Evolved Packet Core (EPC), and ensures user QoE for a particular subscription and service type. Granular control of service quality is critical for operators to establish new business models and monetize services. It enables operators to employ fair-use policies that limit subscriber service abuse (for example, bandwidth hogs such as file sharing), and maintains network performance during peak traffic times.

The policy and charging control is the heart of the Evolved Packet Core, and ensures user QoE for a particular subscription and service type.

Before spending billions of dollars on equipment and deployments, forward-thinking operators will carefully evaluate vendors and proactively measure the QoS and policy management functions of their network devices.

Validating mobile data service quality requires saturating the network with a high load of real-world mobile subscriber traffic, and measuring key performance indicators (KPIs) that identify QoE.

In this paper we will explore the growth in mobile data, and examine how this growth impacts QoE by placing heavy strains on wireless network resources. We will review how wireless networks are evolving from primarily voice-only services to converged voice, video, and data traffic, and how operators can use the emerging 3GPP standards for QoS and policy management to ease network congestion, provide higher service quality, and create a framework for new business models. Finally, we will discuss techniques to validate service quality for mobile broadband and explore a voice over LTE (VoLTE) QoS-validation use case from a tier-one operator.

Policy Management and QoS Considerations for Mobile Broadband

We are beginning an era marked by tremendous global growth in mobile-data subscribers and traffic. Infonetics Research forecasts that mobile data subscribers will grow from 548.9 million in 2010 to 1.8 billion in 2014. HSPA/HSPA+ and LTE show promise in addressing mobile data growth by offering more spectrum capacity and higher data rates at a lower price per bit.

Today's mobile broadband networks carry multiplay services that share radio access and core network resources. In addition to best-effort services, wireless networks must support delay-sensitive, real-time services. Each service has different QoS requirements in terms of packet delay tolerance, acceptable packet loss rates, and required minimum bit rates.

As mobile networks evolve to high-speed, IP-based infrastructure, the wireless industry is ensuring high-quality services by developing QoS and policy-management techniques in addition to adding network capacity. These techniques are designed to ensure application quality, allow operators to offer differentiated services to users, manage network congestion, and recoup the substantial sums that have been invested in building out new networks.

Policy management will play a fundamental role in implementing QoS in mobile broadband. Policy management is the process of applying operator-defined rules for resource allocation and network use. Dynamic policy management sets rules for allocating network resources, and includes policy enforcement processes. Policy enforcement involves service data flow detection and applies QoS rules to individual service data flows. (The following section discusses policy enforcement details.)

Policy management is the process of applying operator-defined rules for resource allocation and network use.

Policy management is critical in three closely-related areas:

- Limiting network congestion
- Enhancing service quality
- Monetizing services

Using QoS and Policy Management to Limit Congestion and Enhance Service Quality

Additional transmission lines, fatter pipes, and improved efficiency are common responses to network congestion. However, this strategy works better for wired networks than for wireless networks. Increasing capacity with additional spectrum and improving spectrum efficiency are important steps in handling the substantial growth of mobile data. However, capacity improvements alone will not solve this complex challenge.

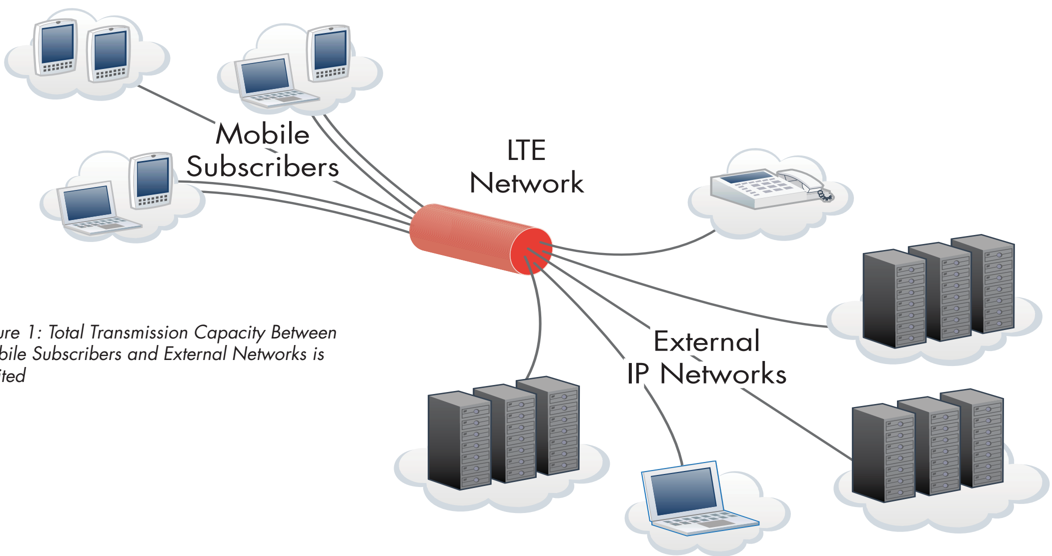

Figure 1: Total Transmission Capacity Between Mobile Subscribers and External Networks is Limited

With proactive management policies, combined with other strategies such as network offloading and demand calibration, mobile broadband networks with finite resources can better-satisfy consumers' demand for multiplay services.

Mobile operators do not have unlimited resources and capital. The radio spectrum is finite, and gains from improved spectrum efficiency can only go so far. Even if operators significantly increase capacity, bandwidth-hungry applications such as peer-to-peer (P2P) services and video will eventually consume any excess capacity. Providing high service quality by over-provisioning network capacity will eventually leave an operator at a competitive disadvantage to providers that offer the same or better QoS, at a lower cost. A solid policy strategy maintains network performance during peak traffic times and spikes in user demand, saving the operator from having to carry excess capacity.

With proactive management policies, combined with other strategies such as network offloading and demand calibration, mobile broadband networks with finite resources can better satisfy consumers' demand for multiplay services. Policy management differentiates services (applications) and subscriber types, and then controls the QoE of each type.

Table 1 demonstrates how subscriber QoE expectation varies by service type. It also highlights how different services have different performance attributes that impact the user's perception of quality. There is a significant distinction between real-time services such as conversational video and voice and best-effort services such as Internet browsing. Real-time services must reserve a minimum amount of guaranteed bandwidth, and are more sensitive to packet loss and latency/jitter.

Services	QoE Expectation	Performance Attributes
Internet	Low – best effort	• Variable bandwidth consumption • Latency and loss tolerant
Enterprise/Business Services	High – critical data	• High bandwidth consumption • Highly sensitive to latency • High security
Peer-To-Peer	Low – best effort	• Very-high bandwidth consumption • Latency and loss tolerant
Voice	High – Low latency and jitter	• Low bandwidth – 21-320 Kbps per call • One-way latency < 150ms • One-way jitter < 30ms
Video	High – low jitter and extremely-low packet loss	• Very-high bandwidth consumption • Very sensitive to packet loss
Gaming and Interactive	Services High – low packet loss	• Variable bandwidth consumption • One-way latency < 150ms • One-way jitter <30ms

Table 1: Comparison of QoE Expectations and Performance Requirements by Service Type

Policy management allows operators to granularly control the availability and QoE of different services. First, policies are used to dynamically allocate network resources – for example, a particular bandwidth can be reserved in the radio base station and core network to support a live video conversation. Next, policy rules control the priority, packet delay, and the acceptable loss of video packets in order for the network to treat the video call in a particular manner.

In other cases, policy rules might be used to limit traffic rates on the network in order to curb network abusers and provide fair use – preventing one user from negatively impacting the quality of another service. P2P file sharing is one example of a very bandwidth-intensive, non-real-time service. P2P services, if left unmanaged, can consume a disproportional amount of network resources and negatively impact the network's ability to establish and maintain real-time service quality.

With the emergence of smart devices (such as smartphones and tablets) the line between who provides value to the subscriber and who they pay has blurred.

Policy Management's Role in New Business Models for Service Monetization

The market landscape is rapidly changing for wireless operators. With voice-only services, operators captured the majority of the customers' mindshare and service revenue. To the consumer, the voice-only wireless operator was viewed as an end-to-end service provider.

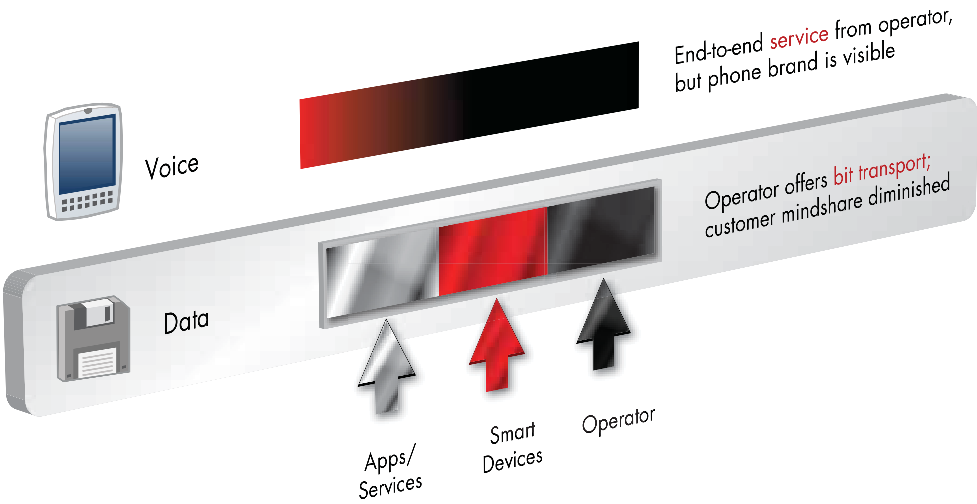

Figure 2: Smart Devices are Diminishing Operators' Revenue by Shifting Consumer Mindshare to Content Providers and Device Manufacturers

With the emergence of smart devices (such as smartphones and tablets), the line between who provides value to the subscriber and who they pay has blurred. Operators are at greater risk of becoming bit transporters, while content/application providers and device manufacturers capture more of the revenue from mobile subscribers. Policy management is one method operators can implement to form new business models and maximize the service monetization.

Policy management helps to retain subscriber mindshare and dollars by allowing granular control of service quality. Policy control enables operators to meet service expectations through network performance modulation, guaranteeing customer QoE and limiting subscriber churn.

Operators will likely take a phased approach in adding policy management to their networks, starting with congestion reduction for applications such as P2P services.

Policy management can also be taken a step further towards the creation of new business models by offering tiered service levels. Tiered service levels can guarantee superior performance and quality to higher paying subscribers (such as corporate accounts). Tiered performance levels can be based on subscription or instant demand. Dynamic policy management allows providers to "put a coin slot in front of the customer." By improving the content delivery quality for fixed periods, policy control supports subscribers' impulse buying of premium services. As an example, a subscriber can upgrade their service for a fixed period of time to watch a video in high definition.

This type of end-to-end network flexibility and service quality control can potentially lead to revenue-sharing agreements with third-party content providers and application content vendors. Operators can form strong relationships with content providers based on excellent service delivery – barring any government regulation preventing tiered service, such as the United States' "net neutrality" policy.

Phasing-In Policy Management

Operators will likely take a phased approach in adding policy management to their networks, starting with congestion reduction for applications such as P2P services. Aggregate-level policy will probably also be introduced in the first phases. It is unlikely that per-subscriber policy management will be implemented early, due to its high complexity. However, as the technology matures, traffic congestion increases, and competitive pressures mount, QoS and policy management will become more and more important. In preparation, operators must make sure they are working with vendors that have a strong framework to supply end-to-end QoS and are capable of supporting evolving needs.

The 3GPP's Vision for QoS and Policy Management in LTE and EPC Networks

The 3GPP's goal is to define an access-agnostic policy control framework, with the objective of standardizing QoS and policy mechanisms for multi-vendor deployments that enable operators to provide service and subscriber differentiation. 3GPP standards explain how to build transmission paths between the user equipment (UE) and the external packet data network (PDN) with well-defined QoS. To this end, the 3GPP has defined an extensive "bearer model" to implement QoS.

Bearer Model

A "bearer" is the basic traffic separation element that enables differential treatment for traffic with differing QoS requirements. Bearers provide a logical, edge-to-edge transmission path with defined QoS between the user equipment (UE) and packet data network gateway (PDN-GW).

Each bearer is associated with a set of QoS parameters that describe the properties of the transport channel, including bit rates, packet delay, packet loss, bit error rate, and scheduling policy in the radio base station.

A bearer has two or four QoS parameters, depending on whether it is a real-time or best-effort service:

- QoS Class Indicator (QCI)
- Allocation and Retention Priority (ARP)
- Guaranteed Bit Rate (GBR) – real-time services only
- Maximum Bit Rate (MBR) – real-time services only

QoS Class Indicator (QCI)

The QCI specifies the treatment of IP packets received on a specific bearer. Packet forwarding of traffic traversing a bearer is handled by each functional node (for example, a PDN-GW or eNodeB). QCI values impact several node-specific parameters, such as link layer configuration, scheduling weights, and queue management.

The 3GPP has defined a series of standardized QCI types, which are summarized in Table 2. For first deployments, a majority of operators will likely start with three basic service classes: voice, control signaling, and best-effort data. In the future, dedicated bearers offering premium services such as high-quality conversational video can be introduced into the network.

For first deployments, a majority of operators will likely start with three basic service classes: voice, control signaling, and best-effort data.

QCI	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100ms	10^{-2}	Conversational voice
2		4	150ms	10^{-3}	Conversational video (live streaming)
3		3	50ms	10^{-3}	Real-time gaming
4		5	300ms	10^{-5}	Non-conversation video (buffered streaming)

Table continued on next page...

Table 2: 3GPP Standardized QCI Attributes

QCI	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
5	Non-GBR	1	100ms	10^{-3}	IMS signaling
6		6	300ms	10^{-5}	Video (buffered streaming) TCP-based (e.g., www, email, chat, FTP P2P file sharing, progressive video, etc.)
7		7	100ms	10^{-5}	Voice, video (live streaming), interactive gaming
8		8	300ms	10^{-3}	Video (buffered streaming) TCP-based (e.g., www, email, chat, FTP P2P file sharing, progressive video, etc.)
9		9	300ms	10^{-5}	

Table 2: 3GPP Standardized QCI Attributes

Allocation and Retention Priority

The 3GPP standards provide mechanisms to drop or downgrade lower-priority bearers in situations where the network become congested. Each bearer has an associated allocation and retention priority (ARP). ARP is used in bearer establishment, and can become a particularly important parameter in handover situations where a mobile subscriber roams to a cell that is heavily congested. The network looks at the ARP when determining if new dedicated bearers can be established through the radio base station.

The 3GPP standards provide mechanisms to drop or downgrade lower-priority bearers in situations where the network become congested.

Guaranteed Bit Rate and Non-GBR Bearers

There are two major types of bearers: guaranteed bit rate and non-guaranteed bit rate. GBR bearers are used for real-time services, such as conversational voice and video. A GBR bearer has a minimum amount of bandwidth that is reserved by the network, and always consumes resources in a radio base station regardless of whether it is used or not. If implemented properly, GBR bearers should not experience packet loss on the radio link or the IP network due to congestion. GBR bearers will also be defined with the lower latency and jitter tolerances that are typically required by real-time services.

Non-GBR bearers, however, do not have specific network bandwidth allocation. Non-GBR bearers are for best-effort services, such as file downloads, email, and Internet browsing. These bearers will experience packet loss when a network is congested. A maximum bit rate for non-GBR bearers is not specified on a per-bearer basis. However, an aggregate maximum bit rate (AMBR) will be specified on a per-subscriber basis for all non-GBR bearers.

Service Data Flows

Service data flows (SDF) are another fundamental concept in the 3GPP's definition of QoS and policy management. SDFs represent the IP packets related to a user service (web browsing, email, etc.). SDFs are bound to specific bearers based on policies defined by the network operator. This binding occurs at the PDN-GW and UE using traffic flow templates (TFT). TFT's contain packet filtering information to identify and map packets to specific bearers. The filters are configurable by the network operator, but at a minimum will contain five parameters, commonly referred to as a 5-tuple. The parameters include:

- The source IP address
- The destination IP address
- The source port number
- The destination port number
- The protocol identification (i.e., TCP or UDP).

The policy and charging enforcement function (PCEF) in the PDN-GW filters packets coming from external networks (i.e., the Internet or corporate VPNs) using TFTs.

Role of Functional Elements in Implementing Policy and QoS

Multiple nodes in the EPC and LTE access play a role in implementing QoS and policy management. The PCRF is the policy server in the EPC. The PCRF takes the available network information and operator-configured policies to create service session-level policy decisions. The decisions, known as PCC rules, are forwarded to the policy and charging enforcement function (PCEF) located in the PDN-GW. The PCEF enforces policy decisions by establishing bearers, mapping service data flows to bearers, and performing traffic policing and shaping.

The PCRF takes the available network information and operator-configured policies to create service session-level policy decisions.

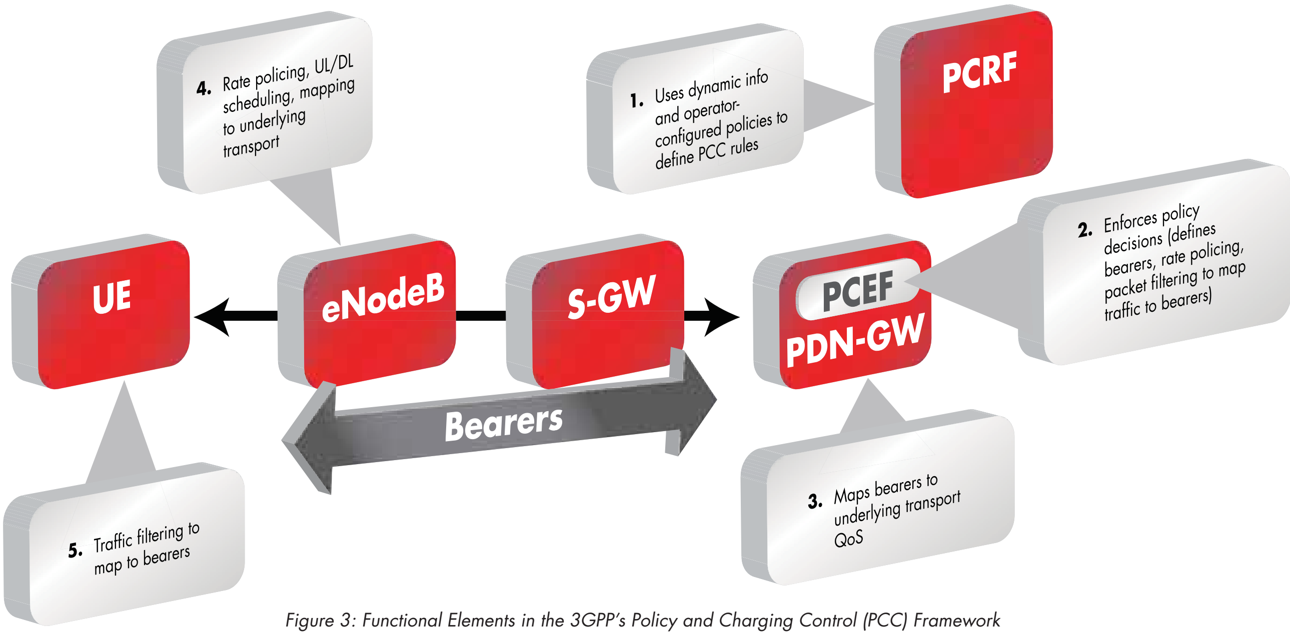

Figure 3: Functional Elements in the 3GPP's Policy and Charging Control (PCC) Framework

The PDN-GW maps bearers to the underlying transport network. The transport network will typically be Ethernet based, and may use MPLS. The transport is not aware of the bearer concept and will use standard IP QoS techniques, such as DiffServ.

The eNodeB is the radio base station in LTE and it plays a critical role in end-to-end QoS and policy enforcement. The eNodeB performs uplink and downlink rate policing, as well as RF radio resource scheduling. It uses ARP when allocating bearer resources. The effectiveness of radio resource scheduling algorithms in eNodeB's has a tremendous impact on service quality and overall network performance. There will be many opportunities for network equipment manufacturers (NEMs) to separate their eNodeB products from other competitive products, and it is something operators must watch closely. The eNodeB, like the PDN-GW, maps bearer traffic to the underlying IP transport network. The UE also plays a role in policy – in the uplink direction, it performs the initial mapping of service data flows to bearers.

Service Quality Validation for Mobile Broadband

Service-quality validation of wireless networks requires saturating the network with a high load of real-world subscriber traffic through mobile subscriber modeling, and by measuring KPIs that identify QoE.

Operators world-wide are spending billions of dollars on equipment, spectrum, and deployments to upgrade their mobile broadband networks. Through network upgrades, operators aim to increase capacities and improve network performance to attract more subscribers, maximize the average revenue per unit (ARPU), and reduce subscriber churn through increased customer satisfaction – at the lowest possible cost. To achieve these goals, operators must carefully evaluate the capacity and performance capabilities of the products they are considering for deployment. After vendor selection, network designs should be prototyped in the lab prior to deployments. As new hardware, firmware, or services are introduced, they should be thoroughly evaluated for performance and it must be verified that they do not negatively impact the performance of existing network services.

Service quality validation allows operators to evaluate networking devices, and proactively measure their QoS and policy management functions. Service-quality validation of wireless networks requires saturating the network with a high load of real-world subscriber traffic through mobile subscriber modeling, and by measuring KPIs that identify QoE. The fundamental strategy is to test the mobile data network with the traffic types and traffic mixes that most-closely resemble the real services that operators will deploy. Service quality and policy/QoS schemes are only stressed when a network encounters congestion. The test approach should involve fully-loading the device or system under test (DUT/SUT). After a network is fully-loaded with a broad mix of real traffic services, detailed QoE measurements are made to quantify network performance. Comprehensive service quality validation arms network equipment manufacturers (NEMs) and operators with a solid method to quickly evaluate the quality and performance of devices and networks.

Mobile Subscriber Modeling

Mobile subscriber modeling is a pillar of any service quality validation strategy. It is the process of defining subscriber types (for example, corporate user vs. casual user), associating applications to a subscriber (such as Internet browsing, email, voice, video, and P2P), and modeling subscribers' usage of applications and their mobility on the network. Subscriber modeling allows testers to replicate real traffic types and usage patterns, and provides the information necessary to fully understand the capacity limits of the network, how multiplay services interact with one another, and the network's ability to differentiate services and subscriber types. Subscriber modeling requires very granular control of service/subscriber emulation.

Figure 4 shows an example of how a causal subscriber might use the network. The subscriber may use the web browser on their smartphone to browse to a web site using a URL. The user pauses to read the web site, and after a certain amount of time clicks on a link to an interesting blog article. While reading the article, the causal user downloads an embedded YouTube video and watches a video clip for 1 minute. Once the video is finished, she might call a friend to discuss the blog and video they just watched.

Figure 4: Functional Elements in the 3GPP's Policy and Charging Control (PCC) Framework

This is a common multiplay scenario carried out by millions of mobile data subscribers every day. Testers need the ability to quickly define specific traffic models in a matter of minutes instead of hours or days. An intelligent application-level user interface is required. The associated mobility protocols for network attachment, security authentication, and bearer establishment must also be emulated. To be effective and emulate a wide and varying degree of application services, the test system must not force the user to be an expert at the underlying protocol procedures.

After modeling the behavior of a specific mobile subscriber, the next step is to place them in a group of like subscribers and model usage over time. This emulates the behavior patterns of different categories of subscribers, such as business users, casual users, and telecommuters.

Traffic usage changes significantly by the time of day (see Figure 5). User behavior patterns should also be mapped to specific times of the day, allowing the emulation of peak usage times. For example, morning service usage is much different than evening traffic mixes.

Subscribers' application-use is rapidly evolving, and the distribution varies greatly by user type. The key point is flexibility. No one can predict exact usage out into the distant future as new applications emerge and become popular. A test frame work must be highly adaptive to future trends.

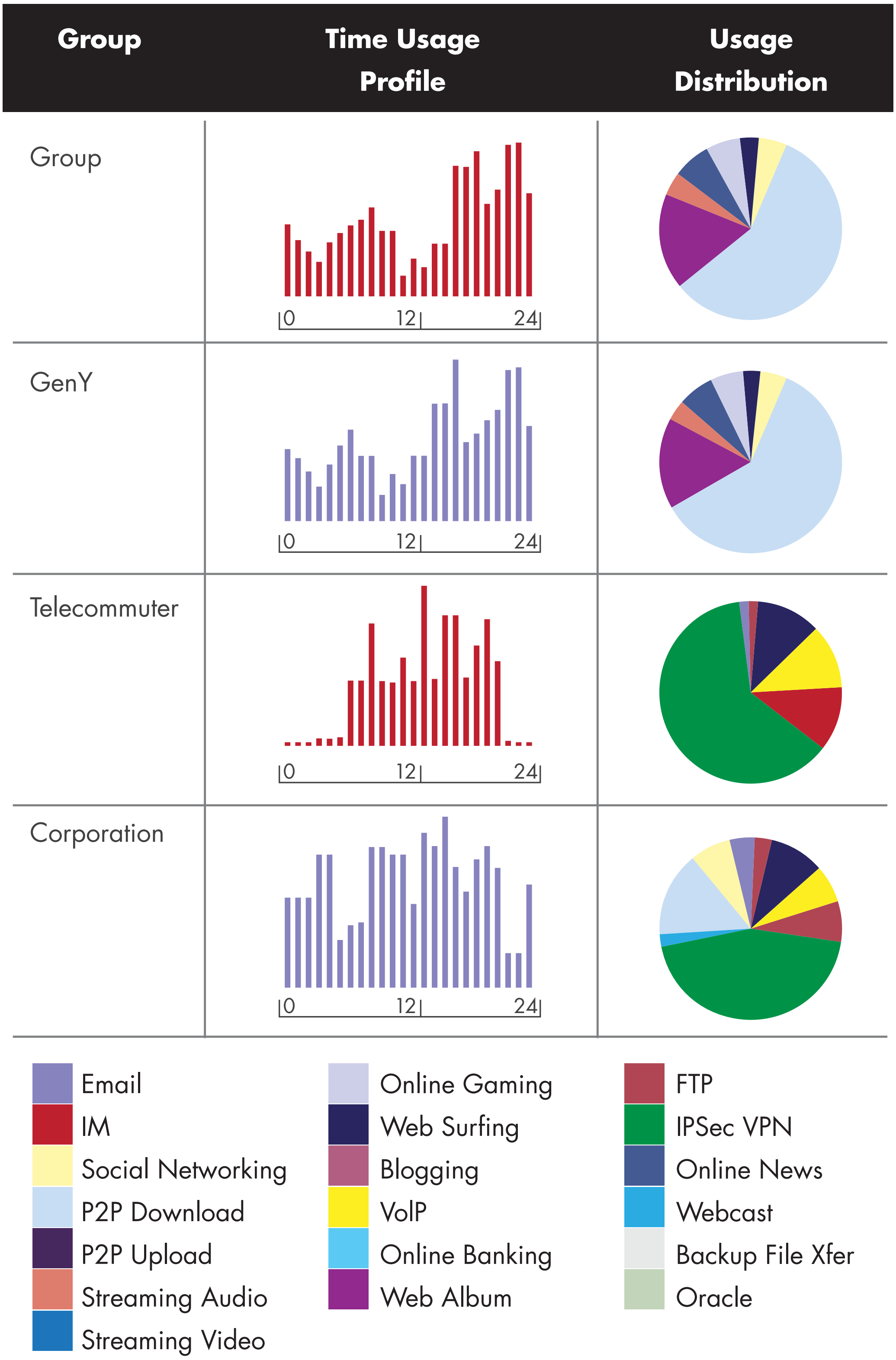

Figure 5: To Attain Real-world Traffic Emulation, Testing Tools Must Model the Specific Behaviors of Each Subscriber Group

Policy and QoS mechanisms must be judged when a network is fully loaded, and there are competing demands for network resources.

Successful testing in the lab using controlled and small volumes of traffic does not guarantee success in the field. Policy/QoS mechanisms must be measured using high volumes of emulated subscriber traffic, when a network or node is at or near capacity. This means generating millions of concurrent web transactions and transactions per second.

The wireless core network is the aggregation point for wireless access network traffic. A core network is called on to terminate the traffic originating from millions of mobile subscribers across thousands of radio base stations. Today, the baseline for testing the wireless core is over 100 Gbps of stateful traffic, and this will evolve to terabits of data as mobile broadband takes off.

KPIs for QoE in Multiplay Wireless Networks

QoE is a measure of the overall level of customer satisfaction with a service. Quantitatively measuring QoE requires an understanding of the KPIs that impact users' perception of quality. KPIs are unique by service type. Each service type, such as conversational video, voice, and Internet browsing, has unique performance indicators that must be independently measured.

Data applications are typically best-effort services, characterized by variable bit rates, and are tolerant to some loss and latency before the user perceives poor quality. Some of the KPIs for data services include:

- Transaction latency (including time-to-first-byte and time-to-last-byte of data)
- Transactions per second
- Concurrent transactions
- Page hits and object hits
- Uplink and downlink throughput
- Re-transmissions
- Failed-transactions

Voice applications are real-time services requiring a constant bit rate. Voice services are sensitive to latency and jitter, but tolerate of some packet loss. The main KPI for voice is the mean opinion score (MOS). MOS_V is a perceptual quality score that considers the effects of codec/quantization level, the impact of IP impairments, and the effectiveness of loss concealment methods.

MOS_V	What does it mean?
5	Excellent
4.5	Very Good
4	Good
3.5	Poor
3	Not Acceptable
2	Severe
1	Useless

Figure 6: Mean opinion score (MOS) scales from 1 to 5 to indicate the transmission quality of video applications over a network

Two important MOS techniques are perceptual evaluation of speech quality (PESQ) and R-Factor. Other important voice KPIs include packet inter-arrival delay (jitter), one-way latency, and the overall connection setup time for a voice call.

Video services have characteristics similar to voice applications. Mobile broadband networks support many different forms of video. Three important categories are live streaming (conversational video), progressive (buffered) download, and adaptive streaming. Live streaming has the highest performance demands. It is a real-time service that is very sensitive to latency, jitter, and packet loss. Perceptual video quality analysis is the most important KPI for video services.

To fully understand QoE, KPIs must be evaluated over time, at varying load rates and application mixes. Policy and QoS mechanisms must be judged when a network is fully loaded, and there are competing demands for network resources. Only under these conditions can the effectiveness of rate limiting/policing, packet shaping, resource scheduling, and packet delay budgets be thoroughly analyzed and tuned.

To properly handle converged multiplay traffic, the network must be capable of giving the operator granular control of QoS/QoE for each type of service.

Use Case: Tier-one operator preparing LTE network to handle voice using VoLTE

One of the world's premier tier-one carriers is equipping its LTE network to handle IP-based voice services. As will most operators, the carrier will first support voice services on LTE handsets using their existing 3G network. Circuit-switched fallback will split voice and data traffic between the 3G and LTE networks, carrying only data on the LTE network. By implementing voice over LTE (VoLTE), the carrier can use its packet-based LTE network for voice service, and more cost-effectively grow its bandwidth and flatten the network architecture to a single IP-based network. But before this deployment can occur, the operator must have high confidence that voice, data, and video traffic can be carried simultaneously with high service quality, ensuring that one service does not negatively impact another.

With actual service turn-up planned in the future, the operator is ramping up its lab. The test lab includes LTE radio base stations, switches, routers, S-GW/PDN-GW, policy server, IP multimedia subsystem (IMS) core network, and other equipment to fully replicate their production network. It also contains Ixia test equipment to "sandbox" different scenarios to validate the network's ability to support VoLTE using IMS-based VoIP. To properly handle converged multiplay traffic, the network must be capable of giving the operator granular control of QoS/QoE for each type of service. This can only be done by properly implementing the 3GPP's standards for policy and charging control.

Objectives

The operator's test objectives are four-fold, emulate different service types, map services to specific bearers, generate high load to create resource contention, and measure the KPIs for each service type:

- Step 1: Emulate user equipment (UE) with multiplay VoIP, streaming video, and data (HTTP/FTP, email, etc.) services and associate them to specific QCIs (bearers). The operator will also configure the network to support different subscriber tiers (enterprise user, consumer, and emergency/government) and speed tiers (2Mbps, 5Mbps, 20Mbps), and map them to QCIs. Since QCIs set the priority and treatment of each traffic type, testing must include the ability to run granular emulations of each type of traffic and associate it to specific QCIs.

- Step 2: Test QCI-to-DSCP (differentiated services code point) mapping to validate proper DiffServ prioritization in the IP transport (backhaul/S1 link). This is important because the Ethernet transport network (switches/routers) does not understand the concept of bearers. Instead, DiffServ is commonly used for QoS in transport networks. Therefore, the eNodeB on one side of the LTE network and the PDN-GW on the other side must map QCIs to DSCP. This mapping between the bearer-aware networks and the transport network must be tested for accuracy.

- Step 3: Put the network under heavy load to create congestion and vary traffic profiles over time. It's only when the network becomes congested and there is competition for resources that we find out if QoS, policy, and prioritization are working properly.

- Step 4: Measure and report KPIs for each QCI (bearer). The operator will define the KPI expectations associated to specific bearers. Thresholds for acceptable packet loss, bit error rate, maximum and average jitter, and voice and video quality scores (MOS) make up the KPIs. The carrier will measure and report KPI's for each QCI over time and with different traffic and subscriber mixes to observe whether they are within the defined threshold limits.

Lab Configuration

The operator's test lab contains a fully-functioning, but scaled-down version of their live network. Figure 7 shows the Ixia-based testbed the operator will use. On the left side of the network there are two distinct ingress points for test traffic. One is Multi-UE emulation directly into the eNodeB over the Uu (RF) interface and the second is eNodeB/MME emulation into the S-GW over the S1-U and S11 interfaces.

To fully saturate the network under test it is necessary to emulate eNodeB/ MME over S1-U/S-11 to provide a high volume of traffic from millions of UEs.

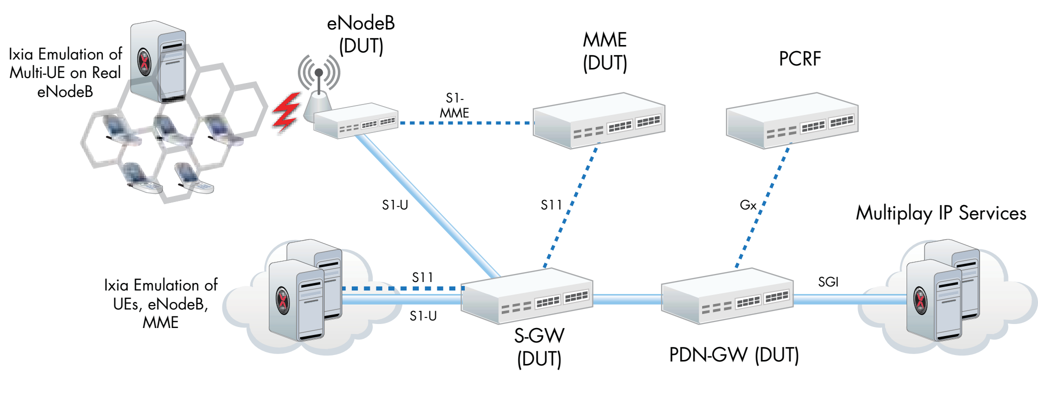

Figure 7: Operator's Ixia-based Testbed for VoLTE

Multi-UE emulation provides end-to-end measurements from Uu to SGi for a select group of UE's. Each emulated UE supports voice, video, and data traffic generation. Dedicated radio bearers are established from the Ixia test system and the service type is appropriately mapped to the correct bearer.

Fully-loading a lab-scale environment with enough traffic to cause congestion and resource contention is sometimes difficult when there are only a handful of base stations available. Even by generating enough traffic to saturate multiple base stations, there is still a limited amount of traffic coming into the packet core network. EPC networks scale to hundreds of gigabits of traffic and millions of subscribers. To fully saturate the network under test it is necessary to emulate eNodeB/MME over S1-U/S-11 to provide a high volume of traffic from millions of UEs. The Ixia test system is also used by the operator to perform these device emulations.

Test Traffic Configuration

The carrier uses settings in Ixia's IxLoad application to configure layer 7 (L7) voice, video, and data activities to specific QCIs and DSCPs. They then measure QCI performance for each of the L7 activities.

For example, KPIs for real-time voice and video applications include:

- Loss packets
- Max and average jitter
- Average latency
- Mean opinion score (MOS)

KPIs for data services (http, ftp, mail, peer-to-peer) include:

- Connection latencies
- Time to first byte of data received
- Time to last byte of data received
- TCP retries
- TCP connection failures

QoS Validation

Now that the environment and test setup is complete, the carrier can measure the performance of the network as they vary input loads, such as traffic rates and types, and subscriber classes (consumer, enterprise, emergency, etc.). In this way, they can vary specific L7 activities and measure the QoS impact on the others.

One of the carrier's many tests is the verification of latency thresholds that ensure delay-sensitive VoIP traffic gets priority over best-effort data traffic. For example, the carrier uses Ixia equipment to emulate a constant level of data traffic (number of subscribers and data rate), while increasing the level of emulated VoIP traffic (see Figure 8 below).

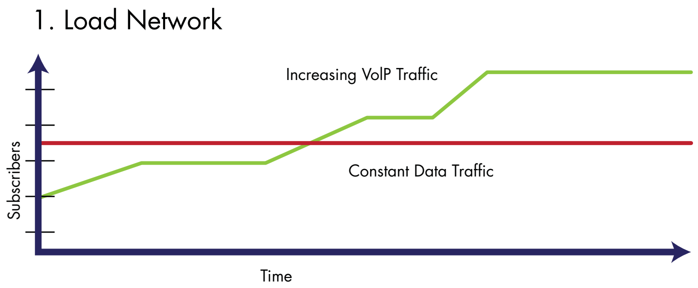

Figure 8: Carrier Test Example, Step 1 —
Load Network with Specific Data and VoIP Traffic

They then use Ixia equipment to measure the network's response to the changing traffic load (see Figure 9). As the total amount of traffic reaches the network's capacity, the carrier's threshold policies should constrain the best-effort data traffic (see red line on graph showing increased latency for data traffic) to free capacity for VoIP traffic (green line, indicating un-changing latency). Since data traffic is delay-tolerant, the carrier ensures satisfactory QoE for both data and VoIP services.

One of the carrier's many tests is the verification of latency thresholds that ensure delay-sensitive VoIP traffic gets priority over best-effort data traffic.

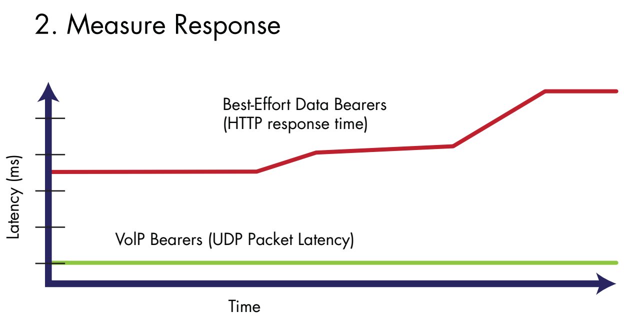

Figure 9: Carrier Test Example, Step 2 —
Measure Response to Ensure Thresholds Produce Satisfactory QoE for Each User Type

The carrier's Ixia-based test system enables them to cost-effectively fine-tune their LTE network and devices to ensure the best possible QoE and to measure and manage the impact of a growing subscriber base and ever-evolving service-use patterns.

Conclusion

The world is at the beginning of the mobile data revolution. The convenience and power of mobile applications delivered on emerging smart devices will fuel the rapid growth in subscribers and sheer volume of data. Operators world-wide are racing to add new services and more powerful devices. They are making substantial investments to upgrade the capacity and performance of their networks. Revenues from voice traffic are relatively flat, and operators will count on new revenue streams from data services to re-coup the money they have invested.

Operators must plan today for the future evolution of the network, which means working with vendors that have a solid roadmap for QoS and policy mechanisms in their products.

If data continues to grow at the rate many analysts forecast, operators will be forced to more-intelligently manage the traffic on their networks. The economic realities and physical limitations of available spectrum prevent operators from simply adding more and more network capacity. To keep their investors happy, operators will be pushed to maximize the revenue from their services. New business models and more premium services will be adopted to achieve this. In turn, as subscribers receive more services that cost more money, their expectations of acceptable network availability and quality will ratchet up. The 3GPP, the world's leading wireless standards body, has had the foresight to plan for these future challenges through their detailed work on QoS and policy management. As the technology and vendor products mature, operators will likely take a phased approach to implementing policy and QoS management, starting with congestion management and evolving into granular control of service quality – ultimately leading to more advanced business models. Operators must plan today for the future evolution of the network, which means working with vendors that have a solid roadmap for QoS and policy mechanisms in their products.

Mobile subscriber modeling and multiplay service emulation are fundamental parts of measuring QoS, policy mechanisms, and QoE. They are mission-critical technique that allow equipment vendors and mobile operators to measure today's network and device performance, as well as adapt to new services and capabilities mobile data networks evolve.

Definitions

2G – Second Generation

3G – Third Generation

3GPP – 3rd Generation Partnership Project

4G – Fourth Generation

AF – Application Function

AMBR – Aggregate Maximum Bit Rate

APN – Access Point Name

APN-AMBR – Access Point Name Aggregate Maximum Bit Rate

ARP – Allocation & Retention Priority

ARPU – Average Revenue per User

BBERF – Bearer Binding & Event Reporting Function

DIFFSERV – Differentiated Services

DL – Downlink

eNB – Evolved NodeB

EPC – Evolved Packet Core, also known as SAE (refers to flatter IP based core network)

EPS – Evolved Packet System (the combination of the EPC/SAE and the LTE/EUTRAN)

EUTRA –Evolved Universal Terrestrial Radio Access

EUTRAN – Evolved Universal Terrestrial Radio Access Network (based on OFDMA)

GBR – Guaranteed Bit Rate

GGSN – Gateway GPRS Support Node

GPRS – General Packet Radio Service

GSM – Global System for Mobile communications

GSMA – GSM Association

HSPA – High Speed Packet Access (HSDPA with HSUPA)

HSPA+ – High Speed Packet Access Plus (also known as HSPA Evolution or Evolved HSPA)

HSS – Home Subscriber Server

IMS – IP Multimedia Subsystem

IP – Internet Protocol

LTE – Long Term Evolution (evolved air interface based on OFDMA)

LTE-A – LTE Advanced

Mbps – Megabits per Second

MBR – Maximum Bit Rate

MOS – Mean Opinion Score

OCS – Online Charging System

OFCS – Offline Charging System

PCEF – Policy Control Enforcement Function

PCC – Policy Charging & Control

PCRF – Policy Control Resource Function

PDN – Packet Data Network

PDSN – Public Data Serving Node

PDP – Policy Decision Point (or Packet Data Protocol)

PDN-GW – Packet Data Network Gateway

PS – Packet Switched

QoE – Quality of Experience

QoS – Quality of Service

QCI –QoS Class Identifier

RAN – Radio Access Network

Rel. 'X' – Release '99, Release 4, Release 5, etc. of 3GPP Standards

RNC – Radio Network Controller

SGSN – Serving GPRS Support Node

SLA – Service Level Agreement

SPR – Subscriber Policy Repository

UE – User Equipment

UL – Uplink

UMTS – Universal Mobile Telecommunications System

UTRA – Universal Terrestrial Radio Access

UTRAN – UMTS Terrestrial Radio Access Network

VoLTE – Voice over LTE

W-CDMA – Wideband CDMA

Ixia Worldwide Headquarters
26601 Agoura Rd.
Calabasas, CA 91302

(Toll Free North America)
1.877.367.4942

(Outside North America)
+1.818.871.1800

(Fax) 818.871.1805
www.ixiacom.com

Other Ixia Contacts
Info: info@ixiacom.com
Investors: ir@ixiacom.com
Public Relations: pr@ixiacom.com
Renewals: renewals@ixiacom.com
Sales: sales@ixiacom.com
Support: support@ixiacom.com
Training: training@ixiacom.com