New Model and Open Tools for Real Testing of QoE in Mobile Broadband Services and The Transport Protocol Impact: The Operator's Approach

Jorge Sandoval, Alfonso Ehijo, Andres Casals, Claudio Estevez
Electrical Engineering Department
University of Chile
Santiago, Chile
jsandova@internetworking.cl

***Abstract*— Real life data from three Mobile Operators has been used to successfully test a New Model of Quality of Experience (QoE) in Broadband Services. The Platform utilized is based on Open Source Tools and it generates Key Quality Indicators (KQI) simulating the final user's perception of the service. The detailed characterization of: a) the consolidated usage profile, b) the transport and application protocols and c) the final service, are the main components of the model that complement the typical network Key Performance Indicators (KPIs). This work also studies how different connection-oriented transport protocols affect the QoE. This provides the operator with an objective comparison of the impact the different protocols have on the QoE metric. The final results are representative KQIs that allow indirect measurements of real life services and service behavior prediction under simulated scenarios. The set of KQIs varies depending upon the market segments and the specific services that each Mobile Operator offers to its customers. From the Regulator's standpoint, when considering several Operators with the same service, this methodology allows the benchmark of QoE.**

Keywords—MBB; Mobile Broadband Service; Mobile Operator; Open Tools; QoE; QoS; TCP; Transport Protocol

I. Introduction

The Mobile broadband (MBB) service is one of the fastest growing in the world, in Chile alone it grew 2 to 4 times in the last three years. In a highly competitive market, operators need to differentiate their service: if the users begin to "perceive a bad experience", they do not recover from it immediately after the improvements are done in the network(s). The key is to know the level of quality of experience perceived by users since "experience" is not merely a KPI. By having an objective metric of the QoE, operators can analyze the weaknesses of the network and hence make the necessary changes.

There are various factors that affect the perception of performance. Some of these factors are: loading times, DNS lookup delay, round-trip times, throughput, etc. To determine an empirical model of the QoE it is necessary to determine which indicators affect the overall performance of the network. As we can see from Fig. 1, the network is mainly composed of TCP-based transmissions. Literature specifies that TCP sessions compose approximately 85-90% [1][2][3] of the Internet traffic; some sources even claim 98% [4]. For these reasons many of the indicators include a study of http and ftp sessions.

Performance indicators are not only, an important contribution of this work, but they are also solutions that increase the studied QoE. As we mentioned earlier TCP-based traffic composed an important percentage of the Internet traffic, so the QoE needs to be taken into account. To increase the QoE, the customer perception of performance, a TCP study is also included. Different TCP algorithms are compared and thus, the impact of the QoE is estimated. All these components will give operators an Internet traffic figure of merit.

II. Background

Performance indicators are important measurements that provide valuable statistical information. These can be used to forecast trends that could serve as an economical prognosis system. The selection of a set of KPIs is essential to provide concrete and valuable information and an objective metric. The set of performance indicators in this study is determined by observing the quality of service parameters that influence most of the perception of the customer, such as throughput and delay. The KPIs used in this work are: session initialization time, DNS lookup delay, round-trip time, HTTP throughput, FTP upstream and downstream throughput, and video streaming interruption events. Not all indicators are weighted in the same way. More on this topic further ahead. One method of increasing the QoE, which is based on these KPIs, is to improve the transport protocol algorithms.

Transport control protocols are a fundamental part of the Internet and can have a great impact on the end-to-end performance of the network. Since it requires feedback from the receiver end, large round-trip times (RTTs) can have an adverse effect on the throughput (and hence on the QoE). Different protocols manage the congestion window size differently, and therefore are less affected by these large RTTs. In a similar way, transport layer protocols may handle loss differently and therefore have different sensitivities to loss, this is why it is good practice to make an assessment on the effect that different transport layer protocols can have on the QoE metric derived.

III. Methodology and Previous Work

In order to understand MBB, the first step is to know which applications the users utilize, which protocols are used by this application, and what factors affect the performance and the experience of their use. Also, theoretical and user traffic analysis is needed to understand the use of this service, how protocols work and other factors. For example, how delay and the parameters of the Operative System affect the performance of the applications in user equipment. Different applications and some of the most used services define the elements to create a QoE model [5] which is then needed to measure and to make a model that teaches and shows the service performance perceived by the users.

IV. Traffic Analysis

Analyzing user traffic allows the operators to know what type of services and applications the users utilize. Web browsing, video and file sharing are among the most used services, but the proportion of each one depends on the region, the country and whether the user is an enterprise or a massive consumer.

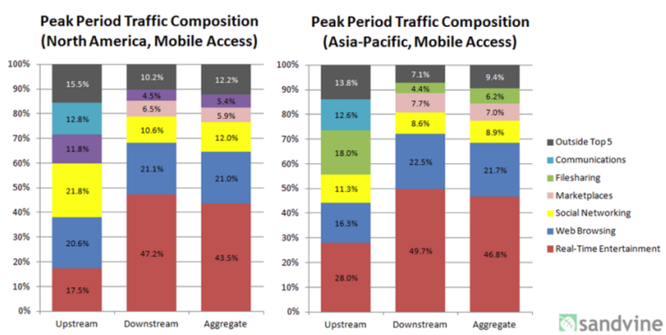

Fig. 1. Traffic Composition in North America and Asia-Pacific.

The figure above shows the composition of traffic in different regions [6], not only can one see different compositions, but one may also see different priorities. This means the most important application for a user in one region may be second in importance in another. It is relevant to know how the most important services work and how, to that end, different classification criteria should be utilized [7]: the protocols utilized, the number of simultaneous connections, Real/Not Real Time protocols, level of user interactivity and volume of data down/uploaded by the application; as shown in table I.

TABLE I. Services Characterisation

Service	Protocols	Simultaneous Connections	RT/NRT	Level of Interactivity	Data Volume
Web Browsing	HTTP, TCP	< 20	RT	Interactive	Low
Network Storage	HTTP, TCP	1 – 20	NRT	Not Interactive	High
P2P	TCP, UDP	~ 100	NRT	Not Interactive	High
Flash Video	HTTP, TCP	1	RT	Interactive	Medium
Social Network	HTTP, TCP	1	RT	Interactive	Low

This characterization allows the operator to understand how this application works. For example, it is important if a user utilizes flash video, if it has a big delay to the server and whether the throughput may not be too high. If the video is in standard resolution this limitation is not a problem but if the video is 3D in Full HD, the application will need a large bandwidth. In the case of Web browsing, the protocol and the application allow multiple and simultaneous connections which enable the use of all available bandwidth.

It is also possible to analyze user mobility by observing how MBB users behave. We can define a criterion of mobility by analyzing what number of cells were visited in a session or day. Mobile, when a user visited a large number of cells; nomad, when this use one cell in the day and another different cell at night, and stationary when user visited only one cell. We can thus conclude that the users are not mobile users (shown in Figure 2).

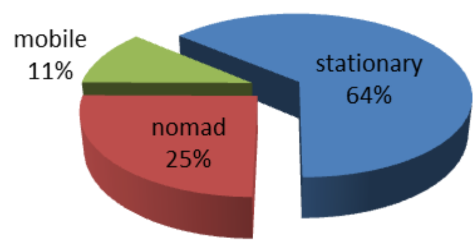

Fig. 2. Mobility of mobile broadband users in Chile

This means that mobile broadband users utilize this service like a fixed broadband. However, the problem arises when the user is in a place with a bad quality internet connection and thus gets bad service all the time because is not in movement.

Another important point is the intensity of use, because in a network, an operator will have different types of users: normal, intensive and heavy. In general, an operator has a few heavy users but when studying different operators in Chile, one may also identify that five percent of all heavy users utilize fifty percent of the traffic; this means that a normal user may have a poor experience when sharing his cell with a heavy user.

V. Theoretical Analysis

According to CAIDA [8], the UDP/TCP ratio in terms of volume is 0.11, this means that 90% of the traffic is over TCP. It is important to understand how this protocol works and how different factors affect the performance of an application. First, it is known that throughput is inversely proportional to delay.

$$Throughput = \frac{RWND}{RTT} \tag{1}$$

When RWND is a reception window of TCP, RTT is a round time trip, and if the user has 300 ms of RTT to the server (considering a typical value of RWND) can one have:

$$\frac{RWND}{RTT} = \frac{2^{16}\,bytes}{300\,ms} \approx 1.6\,Mb/s$$

This means, that the maximum throughput is 1.6 Mbps per connection and if a user is downloading a video flash session, like ©YouTube, independently of an internet access of 10 or 100 Mbps, the maximum throughput remains 1.6 Mbps.

If the operative system on the user's side is RFC1323, this situation may change because the TCP [9] applies a scaling factor (window scale):

$$WS = 2^S \ when \ S \leq 14 \quad (2)$$

In this case, formula (1) of throughput limits change to:

$$Throughput = \frac{RWND \cdot WS}{RTT} \quad (3)$$

And the example in (2), using default values, changes to:

$$\frac{RWND \cdot WS}{RTT} = \frac{2^{16} \cdot 2^4 \ bytes}{300 \, ms} \approx 25.6 \, Mb/s$$

This means, it is conceivable to modify throughput limit per connection only using this RFC. Thus, when one uses an MBB connection one is able to see an improvement of more than 20% in the average throughput by merely activating this option. In the newest operative systems this option is active by default.

Another aspects regards that, if one improves the performance of throughput the delay must be improved as well.

VI. Modeling

It is feasible to build an empirical QoE model by analyzing applications, protocols and the users' behavior. Based on both the traffic and theoretical analysis of MBB users per each one as: service availability, setting up a session timer, throughput, delay, number of interruptions in a video streaming session, DNS lookup time, etc.

The modeling of these indicators is viable when using the following function:

$$QoE = \sum \alpha_i \cdot KPI_i \quad (4)$$

When α_i depends on the importance of KPI_i, it is determined by the observation based on the use of the application shown. Thus, one is possible may to measure the quality of experience indicator reflecting the current state of service, as perceived by the user. The value of α_i can be defined through an iterative process involving feedback surveys of QoE according to ITU-T recommendation [10]. The values for α used in this work are summarized in Table II.

TABLE II. KPI Weights

Service	Weight
Session Initialization Time	5%
DNS Lookup Time	20%
Round-Trip Time	20%
HTTP Throughput	40%
FTP Throughput Downstream	5%
FTP Throughput Upstream	5%
Video streaming interruptions	5%

The values that are dependent on the efficiency of the transport layer design are mainly those related to throughput. Even though RTT has a significant effect on the performance of connection-oriented transport protocols, it is an inherent parameter. This means that the protocol has no effect on RTT, which is why it is not considered for impact computations.

VII. The Connection-Oriented Transport Procotol's Impact on the QoE

The impact of connection-oriented transport protocols on the QoE is an important topic. It is well known that TCP transmission degrades in an inversely proportional ratio to the RTT. This implies that the QoE will depend significantly on the delay between end systems. In this work, the analytical models for the steady-state throughput of four TCP congestion control algorithms are compared: Reno [11] and NewReno [12] (Windows XP), Compound TCP [13] (Windows Vista and higher), and a new protocol called Ethernet-Services Transport Protocol (ESTP) [14]. Even though the protocol is designed to work behind the traffic policing of Carrier Ethernet networks, its QoS configurability is applicable and helpful to the QoE of mobile broadband networks.

ESTP is a QoS-configurable transport protocol. One parameter that can be adjusted is the minimum throughput. If the user has a guaranteed information rate then the slow start and congestion avoidance algorithms do not need to drop below a specific congestion window size. This can be computed in the following manner:

$$cwnd_{\min} = CIR \cdot RTT \quad (5)$$

where the CIR is the committed information rate. The fast recovery/retransmit algorithm decreases the congestion window size as described in the following expression:

$$cwnd_{n+1} = \frac{cwnd_n - cwnd_{\min}}{map(\Delta)} + cwnd_{\min} \quad (6)$$

In this expression Δ is the distance between two packet losses and $map(\Delta)$ is a mapping function that translates this distance into a multiplicative decrease factor. This algorithm guarantees that the throughput will not drop below the agreed rate, which is specified in the service level agreement (SLA). In Fig. 3, it can be observed the traces that correspond to the analytical steady-state throughput models, using $p = 0.001$. As ESTP is configurable, the CIR is set to 20 Mbps. By using a similar method to that of [11], the lower boundary of the analytical model of ESTP is found to be:

$$T = \frac{1}{RTT\left(-\kappa + \sqrt{\kappa^2 + \frac{2\,b\,p\,(1-\beta)}{\alpha\,(1+\beta)}}\right)}$$
$$\kappa = \frac{2\,cwnd_{\min}\,\beta^2 b\,p}{\alpha\,(1+\beta)} \quad (7)$$

The reason this is not the exact throughput is that the mapping function described in [11] is not taken into account. There are two variables that need to be fixed so that the lower boundaries the performance model of ESTP. These values are $\alpha = 1$ and $\beta = 1/2$, which correspond to the additive increase and multiplicative decrease values, respectively.

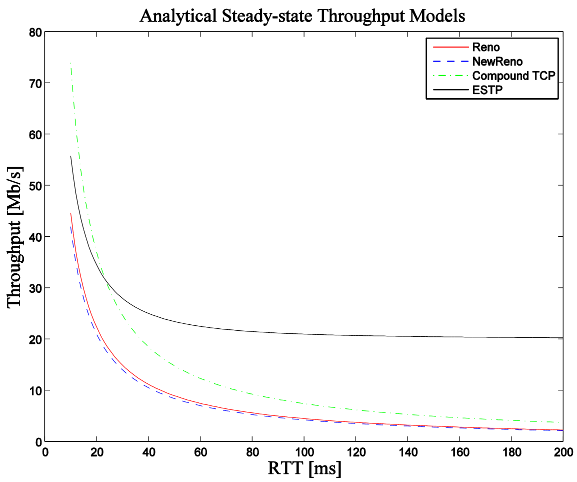

Fig. 3. Steady-state throughputs for Reno, NewReno, Compound TCP and ESTP, using a packet loss rate of 0.001.

Most operating systems (OSs) have their own L4 protocol design, in many cases they are proprietary. This study is limited to open source OSs to facilitate the loading and switching of TCP modules; as well as monitor the associated TCP parameters, such as the receive buffer.

VIII. RESULTS

This model was applied over three Mobile Operators in Chile. In all cases, a QoS platform was used [15] that were adapted in order to measure experience according to the methodology covered in this paper. This platform, built on open source tools, includes two types of monitoring clients: a) a semi-automatic client based on Windows systems and Android smartphone that allows the definition of the exact geographical point on measure, and, b) an automatic client based on Linux that allows locating it on a specific place. A third client behaves as a control point and it's located on the packet core isolating the access problems, as shown in Figure 4.

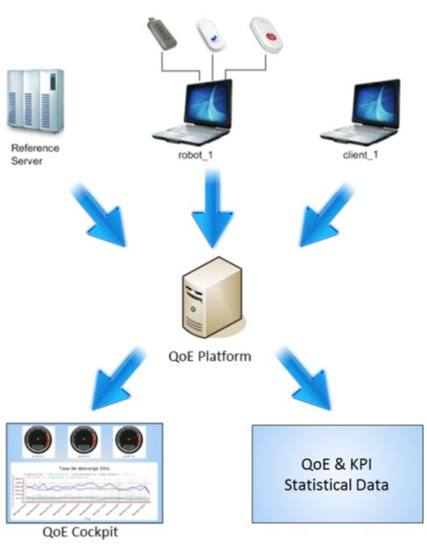

Fig. 4. Monitoring platform architecture.

This Open Source platform allows the benchmarking of QoE and has the location information based on the cell id.

Clients can measure a specific application's performance indicators to rank most used services. For example, top ten web servers, video servers and file sharing servers. Those top servers depend on the region and the county, and even on the specific operator. Using a large number of testers it is possible to have a network indicator using QoE function (4) and KPI weights from TABLE II. , as shown in Figure 5, and thus it is possible to correlate QoE with actual experience perceived by users and see it decreasing in rush hour and when external problems was present.

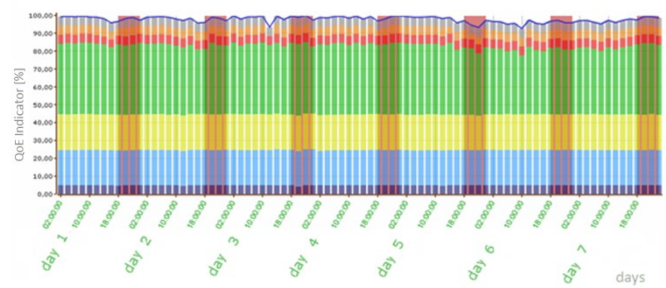

Fig. 5. QoE indicator

To observe the impact on the QoE caused by using more efficient transport protocols it is necessary to compute the throughput improvement of the more efficient protocols over the less efficient ones. It should be mentioned that for the packet loss rate chosen Reno and NewReno perform nearly the same, so the improvement on the QoE is negligible. From Fig. 3, it can be observed that for RTT = 100 ms Compound TCP reaches a throughput performance of 7.4 Mbps and Reno reaches 4.5 Mbps. This corresponds to a 64% improvement. Since three of the indicators depend on TCP-based throughput and these three total 50% of the QoE, it may be inferred that the user will perceive approximately a 32% increase on the quality of experience. For the case of ESTP, the CIR is configurable, so the QoE depends on the CIR. Fig. 6 displays the QoE improvement relative to Reno as a function of CIR for an RTT of 100 ms.

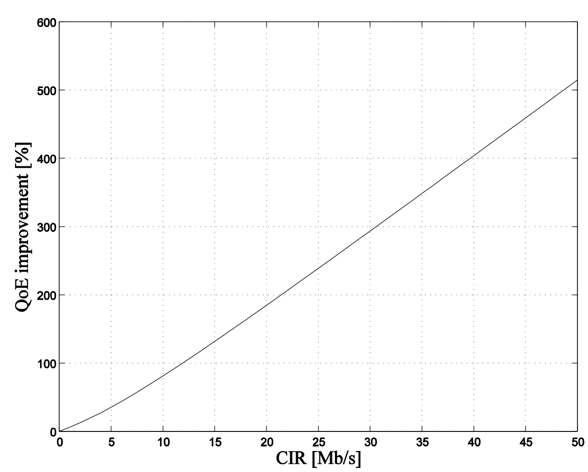

Fig. 6. QoE improvement of ESTP relative to Reno as a function of CIR for an RTT of 100 ms.

The graph portrays the benefit of implementing ESTP on the system. By configuring CIR to the ISP's guaranteed throughput level the user will see an increase of 10% QoE per 1 Mbps of reserved throughput.

IX. CONCLUSIONS

Having a QoE model does not only allow measuring and knowledge of the state of an MBB service, but it also predicts its behavior before major changes such as service downs or degradation.

When we used the model on three different Operators in Chile, it was possible to see that the QoE indicator was highly correlated with the perception of the mobile broadband service.

One important conclusion (and very good news for the LTE fans) is the high impact that the delay in applications has over the quality of experience indicator. Therefore it is important to take actions in order to decrease the delay. For instance, bringing the content close to the client using peering and caching technologies, or improving the access and backhaul technologies.

The impact of transport protocol choice was studied. It was observed that the QoE perceived by the user, as defined in table II, increases significantly. By replacing TCP Reno with Compound TCP there is a 32% increase in the QoE, and by implementing ESTP there is an increase of 10% QoE for every 1 Mbps of CIR that is reserved. This ratio uses a lower bound throughput model, so the improvement could be greater.

ACKNOWLEDGMENTS

This work is partially funded by Proyecto FONDECYT No. 11121655, Fondo Nacional de Desarrollo Científico y Tecnológico, and Program U-INICIA VID 2011, grant U-INICIA 11/14, Vicerrectoría de Investigación y Desarrollo, Universidad de Chile.

REFERENCES

[1] Fraleigh, C., Moon, S., Lyles, B., Cotton, C., Khan, M., Moll, D., et al. (2003, November-December). Packet-Level Tra c Measurements from the Sprint IP Backbone. IEEE Network, 17(6), 6-16.

[2] Williamson, C. (2001, November-December). Internet Traffic Measurement. *IEEE Internet Computing, 5*(6), 70-74.

[3] Thompson, K., Miller, G., & Wilder, R. (1997, November-December). Wide-area Internet Traffic Patterns and Characteristics. IEEE Network, 11(6), 10-23.

[4] Lee, D., Carpenter, B., & Brownlee, N. (2010). Observations of UDP and TCP Ratio and Port Numbers. Fifth International Conference of Internet Monitoring and Protection (ICIMP). Barcelona.

[5] Jorge Sandoval. "Quality of Experience in Mobile Broadband Service". M.Eng Thesis, University of Chile. Santiago Chile, June 2011.

[6] Sandvine. "Global Internet Phenomena Report. 1H 2013". Available in http://www.sandvine.com/downloads/documents/Phenomena_1H_2013/Sandvine_Global_Internet_Phenomena_Report_1H_2013.pdf

[7] Velez, F.J & Correia, L.M. "Mobile broadband services: classification, characterization, and deployment scenarios.". Communications Magazine, IEEE. Volume 40, Issue 4, Page(s):142 – 150. Apr 2002

[8] CAIDA. "Analyzing UDP usage in Internet traffic". May 2009. Available in http://www.caida.org/research/traffic-analysis/tcpudpratio/index.xml

[9] Fall and Stevens. "TCP/IP, The Protocols". Addison-Wesley. Nov 2011.

[10] ITU-T G.1030 Recommendation. "Quality of service and performance – Generic and user related aspects. Estimating end-to-end performance in IP networks for data applications". International Telecommunication Union, Telecommunication Standardization Sector. Nov 2005.

[11] Jitendra Padhye, et al. "Modeling TCP throughput: A simple model and its empirical validation." ACM SIGCOMM Computer Communication Review. Vol. 28. No. 4. ACM, 1998.

[12] Nadim Parvez, Anirban Mahanti, and Carey Williamson. "An analytic throughput model for TCP NewReno." *Networking, IEEE/ACM Transactions on* 18.2 (2010): 448-461.

[13] Kun Tan, Jingmin Song, Qian Zhang, and Murari Sridharan. "A compound TCP approach for high-speed and long distance networks." *International Conference on Computer Communications, IEEE Infocom,* Barcelona, Spain, 2006.

[14] Claudio Estevez, Sergio Angulo, Andres Abujatum, Georgios Ellinas, Cheng Liu, and Gee-Kung Chang. A carrier-ethernet oriented transport protocol with a novel congestion control and qos integration: Analytical, simulated and experimental validation. In Communications (ICC), IEEE International Conference on, pages 2673{2678. IEEE, 2012.

[15] Andrés Casals. "Monitoring Platform for Mobile Broadband Service". Electrical Engineer Thesis, University of Chile. Santiago Chile. Oct 2009.