

The Effect of Router Buffer Size on Subjective Gaming Quality Estimators based on Delay and Jitter

Jose Saldana, Julián Fernández-Navajas, José Ruiz-Mas, Eduardo Viruete Navarro, Luis Casadesus
 Communication Technologies Group (GTC) – Aragon Institute of Engineering Research (I3A)
 Dpt. IEC. Ada Byron Building. CPS Univ. Zaragoza. 50018 Zaragoza, Spain
 e-mail: {jsaldana, navajas, jruiz, evirnete, luis.casadesus}@unizar.es

Abstract—This work presents a study of the effect of router buffer size on the subjective quality experienced by players of online games, which is a service with very tight real-time requirements. Quality estimators for other real-time services like e.g. VoIP were developed years ago, and they are commonly used when planning a new telephony system. They mainly use delay and loss as KPI (Key Performance Indicators). Subjective quality estimators, analogous to VoIP ones, have also been developed for online games, and some of them are based on delay and jitter, but they do not consider packet loss because many games have highly effective loss handling algorithms. This fact has some implications related to the size of router buffers. If it is too big it may add delay and jitter which are not acceptable for gamers. So a study has been conducted, showing that *tiny* buffers (some tens of kB) are more adequate in order to maintain game quality in acceptable levels.

Keywords—online gaming, FPS, subjective quality, MOS, buffer size, QoS, QoE

I. INTRODUCTION

At the beginning, the Internet was designed as a best-effort network, which means delivering information without any delay guarantees. Many years later, it has grown up and is used to provide new services, which sometimes have real-time requirements. One of the first real-time services deployed was VoIP, which is nowadays widely used, and is replacing traditional telephony systems.

The problem of using a best-effort network for deploying a real-time interactive service has been largely discussed, as the users of traditional telephony would not like to change to a new technology unless the offered quality was similar to the one they are used to. Thus, many studies were carried out, trying to identify the different network impairments, and to quantify their effect on the quality perceived by users. One of them was ITU's E-Model [1] which, taking into account delay, packet loss and many other parameters like the codec used, provides a quality estimator without having to repeat the surveys each time a new VoIP system is deployed.

In the last years some new services have arisen, and among them online gaming has become one of the most popular. Online games can be divided into a number of genres. In this paper we will study the ones which have the most stringent real-time requirements: FPS (First Person Shooters), which create a virtual scenario shared by some tens of users who have to accomplish a mission or kill all the enemies. These

games use UDP, because the game has high interactivity since the player has direct control of the avatar.

Some studies [2], [3] have made an effort in order to deploy quality estimators similar to the E-Model, but adapted to a specific game. As we will see in the next section, the first step they take is to identify the network parameters that impair the experienced quality, named KPI (Key Performance Indicators), which are then properly combined so as to obtain a MOS (Mean Opinion Score) formula. As we will see, instead of considering delay and packet loss as the main KPI, these studies use delay and jitter.

This fact is very interesting and has some implications related to the scenario where players are, i.e. they are usually connected to the Internet via an access network with a limited bandwidth and a bottom or mid-range router. These routers can have different behaviours depending on access bandwidth and the implementation of the output buffer. In this paper we study the effect of the size of this buffer when using quality estimators based on delay and jitter.

The paper is organized as follows: section II discusses the related works about buffer sizing and quality estimators for online games. The next section covers the tests and results, and the paper ends with the conclusions.

II. RELATED WORKS

A. The buffer sizing problem

In the last years, many studies related to the buffer size problem have been published. A very complete review of the works related to the topic can be found in [4]. The problem is normally considered for backbone routers and TCP flows. As a summary, we can say that the traditional *rule of thumb* used to calculate buffer size was the bandwidth-delay product (BDP). In 2004, Appenzeller et al. deployed the so-called *Stanford model* [5], which proposed the use of *small* buffers, calculating the size as the quotient of BDP and the square root of the number of TCP flows. Other model is the *tiny* buffer [6], which considers that a capacity of some dozens of packets is enough so as to obtain an utilisation about 80-90%. Finally, some works [7], [8] consider the combined effect of TCP and UDP on *tiny* buffers, and an anomalous region where packet loss grows with buffer size was found.

The present work is centered on the impairments for real-time traffic when it shares a buffer with other traffics. We will

not consider backbone routers, as internet access is usually provided by smaller and simpler ones. Different buffer sizes will be tested and compared in order to have an idea of the influence of buffer size on the performance of real-time services. We will consider drop-tail FIFO *tiny* buffers of some tens of kB, as bigger ones would increase the delays, making the experienced quality of real-time services unacceptable.

B. Measuring subjective quality for online games

The problem of obtaining estimators for the subjective quality of games has been issued in many works. In [9] only the delay was considered, and it was denoted as SRT (System Response Time), which was defined as the time necessary to detect a user event, to process it, and to display the new game state at the output device of the user. A simple MOS was then calculated, which had a linear dependence on the SRT.

Another study [10] carried out a set of experiments, using an emulator to add controlled delays and packet losses to the network, and asking real players to fill in some questionnaires regarding to their perceived quality. Two FPS games were tested: *Halo* and *Quake III*. The work studied separately the effect of delay and loss, so a MOS formula was not developed. Some of the conclusions of this work were that delay has a bigger influence than loss. Another interesting result related to packet loss was found: while *Halo* did not work if 4% of the packets were lost, *Quake III* could work even with a 35% of packet loss. This means that some games have methods to compensate different KPI. The study did not consider the jitter as a KPI, as they considered that its effect on perceived quality was significantly smaller than the latency impairment.

In [11] a survey was conducted and also a practical evaluation of four games. Only delay and jitter were considered as KPI, while packet loss was left for future work. The results showed very different MOS impairments of the KPI for each game. The first MOS model for a FPS (*Quake IV*), adapted from the VoIP E-Model, was presented in [2], showing that this game, like its predecessor, has a very effective loss handling algorithm, so the two KPI considered were delay and jitter. A polynomial formula for the MOS was obtained using a multi-dimensional regression analysis. This is the model we are going to use in the current work.

A similar analysis was carried out in [3] and a formula for the MOS of an MMORPG (Massive Multiplayer Online Role Playing Games) was developed (*World of Warcraft*). Different combinations of loss and jitter were added using an emulator, and a group of users had to play and fill in some questionnaires regarding to the perceived quality.

III. TESTS AND RESULTS

A. Scenarios of Interest

The scenario we are going to consider can be seen in Fig. 1. A number of hosts are connected via the same access router to a server. If we consider gaming service, we can find this scenario in an Internet café. Gaming has been reported as one of the main activities deployed by the users of these

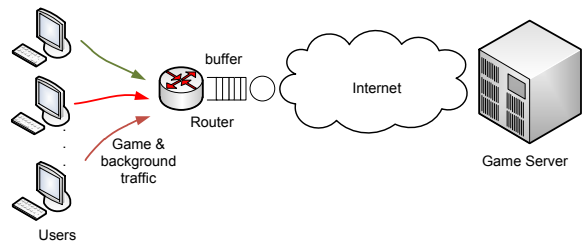


Figure 1. Scenario considered.

businesses, which are still a very popular way for connecting to the Internet. This scenario was more deeply studied in [12].

In this scenario, the connection will be shared by many services, some of them real-time. The total traffic offered to the router may vary and in some moments it can even be over the link capacity. So the router will add delays and discard packets, and the buffer size and policy will have an influence on the distribution of these impairments. In particular, some policies are packet-size aware, as they penalize different sizes in different manners, as we will see later.

B. Previous results using VoIP

In previous works, our group studied the influence of the buffer size on other real-time service, i.e. VoIP traffic [13], and the results showed that the MOS has a monotonically decreasing behavior as the background traffic grows. In that case, jitter was not considered as a KPI, as a de-jitter buffer is used, which adds some delay and discards packets that arrive out of time. This can be done as the codec and consequently the packet rate, are known. In that study the VoIP native traffic showed a good behavior when using a small buffer, as small packets have less probability of being discarded than big ones. Acceptable MOS results were obtained even with an offered traffic over the bandwidth limit. In VoIP the MOS is considered acceptable over 3.5.

C. Tests methodology

When considering online gaming, as the application is not open, we do not know whether it uses a de-jitter buffer or not. As we have seen, delay and jitter are the KPI considered to affect certain games, so in this study we will measure their effect. We will consider the same scenario used for VoIP service (Fig.1), i.e. we have a number of players sharing the access network.

In order to test the most stringent real-time applications, we have selected a FPS game for which a MOS model exists [2]: *Quake IV*. First, we needed traffic traces of the game under study. We obtained them from the CAIA project [14], where many online games have been analyzed, and some traffic traces are available. The traces were obtained from real parties in controlled conditions, and are very well documented. They contain a number of packets which is the product of 5,000 and the number of players. Fig. 2 shows the histograms for packet size and inter-packet time for the selected game. The average size of the packets at IP level is 79.5 bytes. The game generates 64 packets per second. So the total bandwidth generated by each player is roughly 40.7 kbps.

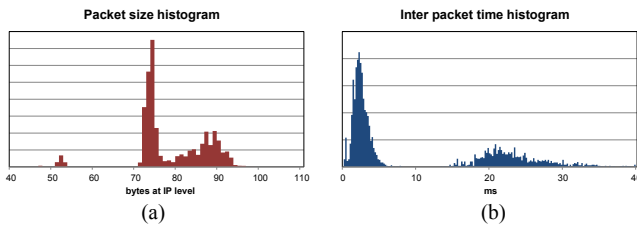


Figure 2. a) Packet size and b) inter-packet time histograms for *Quake IV*.

We will study the most stringent link, which in our case is the uplink, as many access technologies are asymmetric and the bandwidth of the uplink is significantly smaller than the one for the downlink. We have first separated the client-to-server traffic and then combined it in order to obtain different numbers of users, as done in [12].

In this case, we will consider 20 players sharing an Internet access with uplink bandwidths of 2 Mbps and 3 Mbps, and a drop-tail FIFO buffer (Fig. 1) with four possible sizes: 10 kB, 20 kB, 50 kB and 100 kB. They can be considered as *tiny* buffers [4]. For each point, 400 seconds of the application traffic have been sent, sharing the buffer with different amounts of background traffic, which ranges from 0 to the traffic limit of the connection. A background traffic distribution with Poisson inter-packet times and three packet sizes is used: 50% of the packets were 40 bytes at IP level, 10% had 576 bytes and the rest were of 1,500 bytes [15].

A Matlab-based simulator has been used in order to introduce the obtained game traces and the background traffic

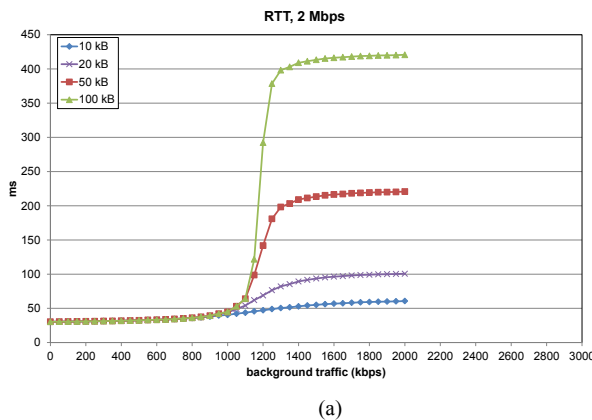
in the same drop-tail buffer. Delay and jitter are measured round trip. The delay is composed by the buffer delay plus an additional RTT of 30 ms with a variance of 5, which can be a typical network delay for a regional scenario [16], and the response time of the game server. The acceptable delay boundary has been said to be around 200 or 225 ms [10].

D. Test results

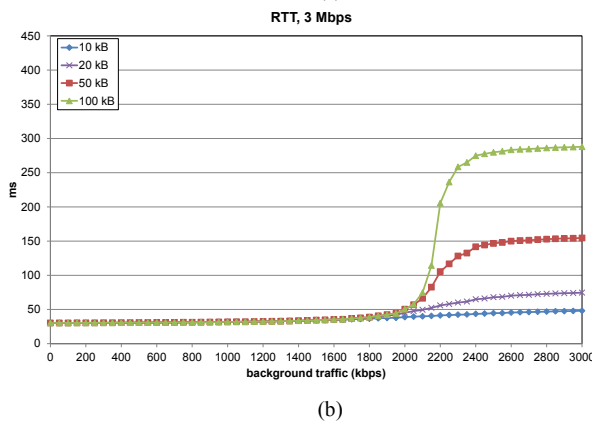
Fig. 3 a) shows the total RTT for the 2 Mbps connection. On behalf of clarity, we have represented both 2 Mbps and 3 Mbps buffers with the same scale for the x-axis. It can be seen that, when bandwidth limit is reached, which happens at 1200 kbps, the delay grows up depending on the buffer size. In this case, the smaller the buffer, the smaller the delay. We can observe that the 100 kB buffer cannot be used if the total offered traffic is above the bandwidth limit. The buffer of 50 kB is in the limit of acceptable values (220 ms), which means that the buffer should be smaller than this value.

Fig. 3 b) has a similar behaviour, growing up when the bandwidth limit is reached. But in this case the bigger bandwidth limit makes the delay become smaller. So in this case, with a buffer of 50 kB we will obtain acceptable delay values (around 150 ms).

The used quality model [2] considers jitter as the standard deviation of the delay. Jitter is measured in different ways in the literature, and there is not a clear consensus on its definition. In VoIP it can be measured as inter-packet delay

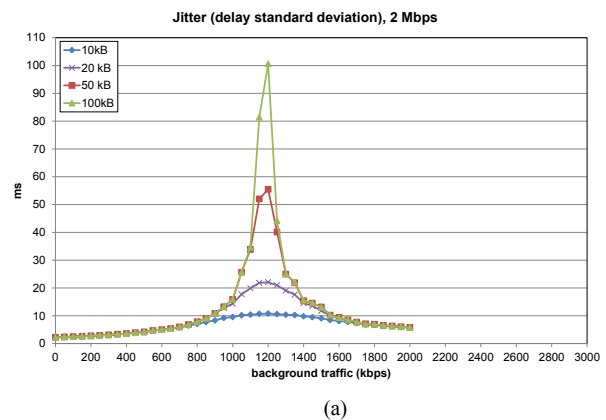


(a)

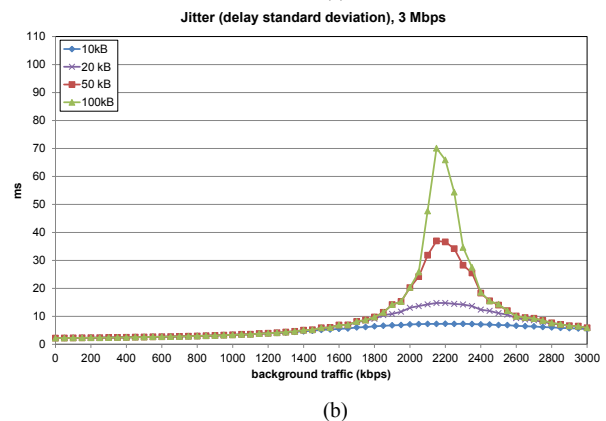


(b)

Figure 3. RTT as a function of background traffic) 2Mbps. b) 3 Mbps



(a)



(b)

Figure 4. Jitter as a function of background traffic: a) 2Mbps. b) 3 Mbps

variation but, in the case of the considered game, it cannot be measured that way, as inter-packet time is not fixed (Fig. 2b).

Fig. 4 a) shows the jitter for the 2 Mbps connection. For the smallest buffer, it shows a good behaviour but, when bigger buffers are used, a peak appears when the offered traffic is roughly the bandwidth limit. This is caused by the saturation of the access link, which makes the buffer occupation grow. But if the offered traffic is over the bandwidth limit, then the jitter gets reduced, because the buffer is always full, so all the packets will experience the same delay.

If we compare Fig. 4 a) and b), we can see that the peak for the 100 kB graph has been reduced from 100 ms to 70 ms, and this decrease can be observed for the rest of buffer sizes. This means that there is a relationship between the bandwidth limit and the maximum buffer size allowed: the bigger the bandwidth of the access, the bigger the buffer allowed.

Fig. 5 shows the packet loss for each packet size, using the 10 kB buffer, where it can be seen that small packets have a clear advantage when drop-tail buffer policy is used. It should also be noticed that for a fixed buffer size, the amount of packet loss is smaller when the bandwidth limit is bigger.

We will finally present the results of the perceived quality. The formula proposed in [2] first calculates a *network impairment* parameter named X , which depends on RTT (*ping-average*) and *jitter_average*:

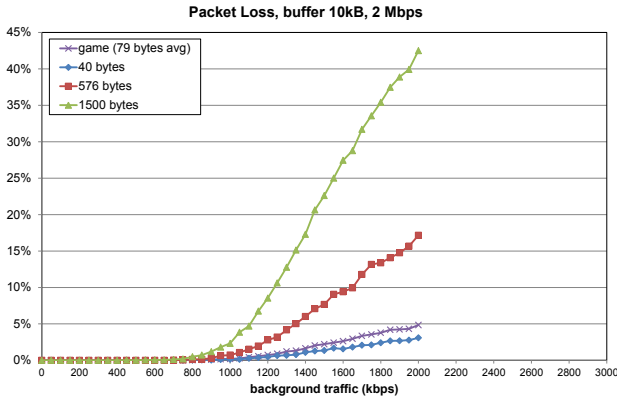
$$X = 0.104 * ping_average + jitter_average \quad (1)$$

The MOS formula depends on X :

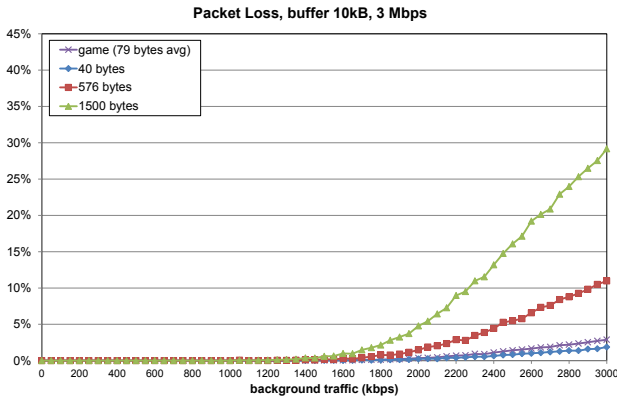
$$MOS = -0.00000587X^3 + 0.00139X^2 - 0.114X + 4.37 \quad (2)$$

As said in Section II, packet loss under a threshold is not considered to significantly affect the quality, as this game has a packet loss concealment algorithm. This also happened in *Quake III*, which surprisingly, could work properly even with packet loss near to 35% [10]. In the current study packet loss for game traffic is always below that percentage (Fig. 5). Although some studies use the VoIP scale and consider MOS acceptable when it is above 3.5, some others consider that a value of 3 can be good enough, and that gamers will try to find another server when MOS is about 2 [11].

Fig. 6 a) and b) present the MOS. These graphs show a surprising behaviour when compared with the MOS of VoIP, which has a monotonically decreasing behaviour. When the background traffic is small, the behaviour is as it could be expected: the bigger the background traffic, the worse the experienced quality. When we approach to bandwidth limit, the graphs become worse, due to the jitter increase. But surprisingly, when the offered traffic is above the bandwidth limit, the experienced quality grows up as the background traffic increases, achieving better values. This anomalous

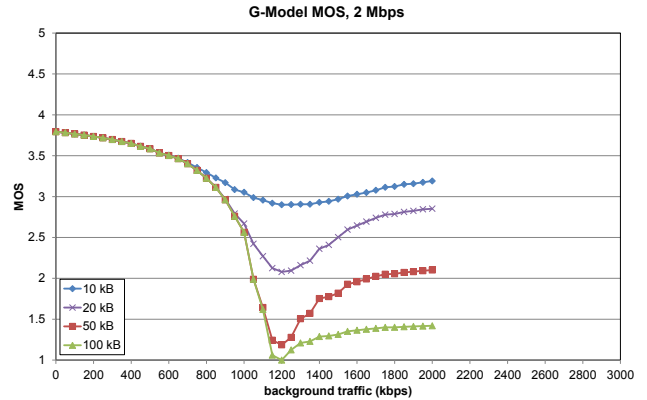


(a)

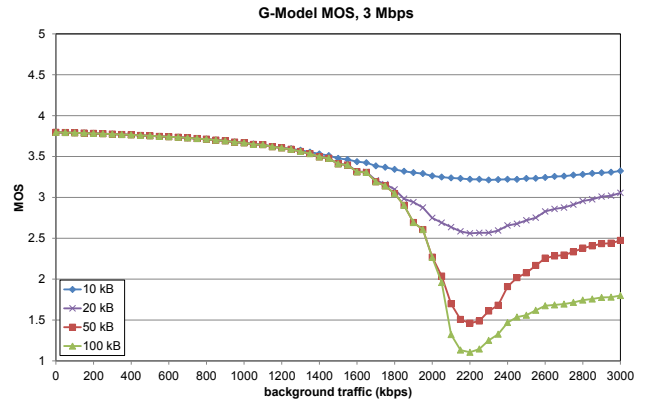


(b)

Figure 5. Packet loss for each traffic for 10kB buffer



(a)



(b)

Figure 6. MOS as a function of background traffic: a) 2Mbps. b) 3 Mbps

“valley” behaviour is caused by the jitter reduction shown in Fig. 4. As we have said, the jitter is significantly reduced because the buffer is always full, so the delay is roughly the same for every packet.

If we consider acceptable the MOS values around 3, in both figures it can be seen that only the smallest buffer (10 kB) achieves acceptable MOS results despite the offered traffic. The rest of buffer sizes maintain acceptable MOS levels until the total offered traffic is around 90% of the bandwidth limit. But for the 20 kB buffer with 3 Mbps bandwidth, if the offered traffic grows above the bandwidth limit, acceptable MOS values can be reached again. This effect is caused by the decrease of the jitter.

E. Discussion of the results

The delay results are not surprising: as the buffer gets smaller and the bandwidth gets bigger, the buffer can be emptied more quickly. But the jitter presents a peak, and then it is reduced because the buffer is always full. This causes the surprising behaviour of the MOS graphs. Although the results have been obtained using a concrete quality model, this behaviour can be expected to be similar for other subjective quality estimators based on delay and jitter.

On the other hand, packet loss does not affect the game traffic. With the used buffer policy, it has been measured to be small, as drop-tail FIFO policy penalizes big packets, and the ones of the game are small.

To sum up the results, we can highlight the importance of integrating all the parameters into a MOS value. We cannot simplify the problem studying each of the KPI separately. In the case of the application under study, they are delay and jitter, which are affected by the bandwidth of the access, as expected. But the results show that the buffer size has a strong influence too. It has also been shown that a bigger bandwidth permits bigger output buffers. An application which runs well in a local environment may experience problems when using an access network to interact with a game server located on the Internet, so buffer size is a crucial parameter which has to be configured taking into account all these relationships, especially for commercial access routers.

IV. CONCLUSIONS

This work has studied the effect of the buffer size on the perceived quality for certain online FPS games. Subjective tests have shown that their main KPI are delay and jitter. Packet loss is not considered unless its amount gets very high.

Simulations have been conducted in order to obtain the delay and jitter produced by the buffer of the access router, using different amounts of background traffic. Network delay and jitter have also been added. The traces of the game were obtained from real parties and were properly combined in order to obtain the traffic of 20 simultaneous players sharing the same Internet access. The results show a jitter peak that causes a “valley” in the MOS graph, obtaining an anomalous region in which the MOS grows with the background traffic.

Small buffers present better results than bigger ones. It would be interesting for game developers to identify the behaviour of the router buffer so as to adapt the traffic to it. The relationship between buffer size and connection bandwidth has to be taken into account.

Future works may study this relationship when other buffer policies different than drop-tail FIFO are used, showing the benefits of buffers that prioritize real-time traffic.

ACKNOWLEDGMENT

This work has been partially financed by CPUFLIPI Project (MICINN TIN2010-17298), MBACToIP Project, of Aragon I+D Agency and Ibercaja Obra Social, Project of Catedra Telefonica, Univ. of Zaragoza.

REFERENCES

- [1] “The E-model, a computational model for use in transmission planning”, ITU-T Recommendation G.107, 2005.
- [2] A. F. Wattimena, R. E. Kooij, J. M. van Vugt, O. K. Ahmed, “Predicting the perceived quality of a first person shooter: the Quake IV G-model”. In Proc. 5th SIGCOMM workshop Network and system support for games (NetGames '06), ACM, New York, NY, USA, 2006.
- [3] M. Ries, P. Svoboda, M. Rupp, “Empirical study of subjective quality for Massive Multiplayer Games”, Systems, Signals and Image Processing, 2008. IWSSIP 2008. 15th International Conference on, pp.181-184, 25-28 June 2008.
- [4] A. Vishwanath, V. Sivaraman, and M. Thottan. “Perspectives on router buffer sizing: recent results and open problems”, SIGCOMM Comput. Commun. Rev. 39, 2 (March 2009), 34-39.
- [5] G. Appenzeller, I. Keslassy, and N. McKeown. “Sizing router buffers”, In SIGCOMM '04, pages 281–292, New York, USA, 2004. ACM Press.
- [6] M. Enachescu, Y. Ganjali, A. Goel, N. McKeown, T. Roughgarden. “Part III: routers with very small buffers”. SIGCOMM Comput. Commun. Rev. 35, 3, pp 83-90, Jul 2005.
- [7] A. Vishwanath and V. Sivaraman. “Routers with Very Small Buffers: Anomalous Loss Performance for Mixed Real-Time and TCP Traffic”. In IEEE IWQoS, Netherlands, 2008.
- [8] A. Vishwanath, V. Sivaraman, G. N. Rouskas. “Considerations for Sizing Buffers in Optical Packet Switched Networks”. In IEEE INFOCOM, Brazil, 2009.
- [9] C. Schaefer, T. Enderes, H. Ritter, M. Zitterbart. “Subjective quality assessment for multiplayer real-time games”. In Proc 1st workshop on Network and system support for games (NetGames '02). ACM, New York, NY, USA, 74-78. 2002.
- [10] S. Zander, G. Armitage, “Empirically Measuring the QoS Sensitivity of Interactive Online Game Players”. In Proc. Australian Telecommunications Networks & Applications Conference (ATNAC 2004), Sydney, Australia, Dec. 2004.
- [11] M. Dick, O. Wellnitz, L. Wolf. “Analysis of factors affecting players' performance and perception in multiplayer games”. In Proc. 4th ACM SIGCOMM workshop on Network and system support for games (NetGames '05). ACM, New York, NY, USA, 1-7, 2005.
- [12] J. Saldana, J. Murillo, J. Fernandez-Navajas, J. Ruiz-Mas, J. I. Aznar, E. Viruete, “Bandwidth Efficiency Improvement for Online Games by the use of Tunneling, Compressing and Multiplexing Techniques,”. Proc. Int. Symp. Performance Evaluation of Computer and Telecom. Systems SPECTS 2011, pp.227-234, The Hague, Netherlands, June 2011.
- [13] J. Saldana, J. Murillo, J. Fernandez-Navajas, J. Ruiz-Mas, E. Viruete Navarro, J. I. Aznar, “QoS and Admission Probability Study for a SIP-Based Central Managed IP Telephony System,” New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference on, Feb. 2011.
- [14] L. Stewart, P. Branch, Quake4, Map: q4dm1, 7players, 20Jul2006. Centre for Advanced Internet Architectures SONG Database, http://caia.swin.edu.au/sitrc/song/files/quake4_310706_1_q4dm1_7_fragment.tar.gz
- [15] Cooperative Association for Internet Data Analysis, “NASA Ames Internet Exchange Packet Length Distributions”.
- [16] AT&T Global IP Network, <http://ipnetwork.bgtmo.ip.att.net/pws>