Video Quality as a Driver for Traffic Management with Multiple Subscriber Classes

Martín Varela and Janne Seppänen
VTT Technical Research Centre of Finland
{martin.varela, janne.seppanen}@vtt.fi

1. Introduction

Video services currently account for a very large portion of the total traffic on the Internet, and this portion is foreseen to keep rising [*1*]. This trend, coupled with the resource-hungry nature of video services, poses significant problems for network management, if good perceptual quality levels are to be achieved. In mobile networks, in particular, this can be a problem when a cell contains several users streaming video concurrently. In this paper we present a short overview of a multi-faceted mechanism for cross-layer quality-driven traffic management for video services at the last hop, which we have proposed in [*2*]. We consider over-the-top (OTT) services, where the network operator does neither control the content nor profit directly from it. Despite ever-more-efficient encoding schemes, mobile video traffic is poised to keep increasing its need for resources, as high-resolution displays appear in mobile devices, and users become accustomed to HD video on their TV and desktop/laptop systems. Since bad quality might lead to user churn, solutions in the form of access control, or *Differentiated Services*, have been explored, which may allow implementing network QoS mechanisms that result in better QoE for the end users. An immediate problem that appears in this context is that of identifying the traffic to mark as high-priority. In the case of RTP-based streams, simply looking at packet streams might be sufficient, but with the majority of OTT services being HTTP-based, the problem becomes non-trivial.

Research on quality-driven traffic management for video services has been done for IPTV (e.g. in [*3*] [*4*]), and to a lesser extent on OTT services [*5*] [*6*] in wireless contexts.

2. A multi-faceted approach

In this work, we propose a composite approach to managing the traffic in order to provide adequate QoE to the users. We propose a subscriber-based differentiation scheme (implemented, without loss of generality, with two classes of users, namely *premium* and *normal*), and a traffic management scheme based on both access control and application-based traffic differentiation.

Overall, our solution works as follows. New flows arriving at a generic *Access Point* are classified both by their subscriber class and by their application type (the latter classification is done by using the two-stage statistical classifier described in [*7*]). Regardless of subscriber class, inelastic flows are only admitted if the average Mean Opinion Score (MOS) of other video streams is above a set threshold (note that premium users cannot preempt normal users, and so a premium user's stream might be dropped even if normal users are currently streaming video). If a flow is admitted, then it is assigned to a queue with the adequate priority based on its application type, subscriber class, and current estimated QoE. The objective of this process is to ensure that a) admitted video streams provide acceptable quality, b) premium users' streams achieve better quality when congestion arises, and c) the system is fair to normal users as well (not preempting them, and interleaving the priority of premium and normal users' application classes).

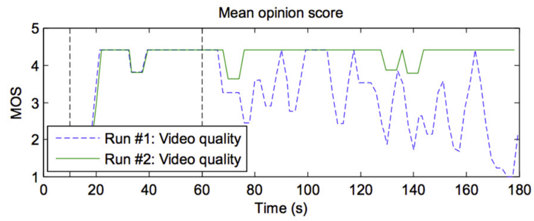
All flows enter a single FIFO queue, and individual flows are promoted on an as-needed basis to higher priority queues depending on their current quality, subscriber class, and application class. In particular, a threshold of 3.0 in the usual 5-point MOS DC scale is set, so that actions are taken when a flow's quality estimation drops below this value. A hysteresis mechanism is then implemented, ensuring that the improved quality achieved by promoting the flow is stable for a set period of time over a second threshold (4.0 points in the case of the reported results) before returning the flow to a lower-priority class (if possible). Queues of higher priority are emptied before those of lower priority (up to a certain limit), and traffic within each queue is handled with stochastic fairness queuing [*8*].

The quality estimations were performed using a PSQA [*9*] based model for IPTV-like services. The traffic management was implemented on top of a Linux-based router, by using the Hierarchical Token Bucket (HTB) queuing discipline [*10*]. HTB is rather complex, but provides a very flexible approach to handling different traffic classes. The system uses both the priority of a class and a set limit for each class to achieve fairness. Within each class' queue, stochastic fairness queuing (as implemented in Linux [*11*]) is used.

3. Performance evaluation

The performance of the proposed approach was tested as a proof-of-concept in a laboratory environment. Four different aspects of the proposed system's performance were tested, namely 1) responsiveness in the case of congestion (application differentiation), 2) subscriber priority handling, 3) reaction times, and 4) admission control. All tests were done between 15 and 40 times, and the results presented herein are representative of the average behavior of the system.

In the first set of tests, a test video stream was subjected to contention by a large bulk transfer. Figure 1 shows the results of the first test set.



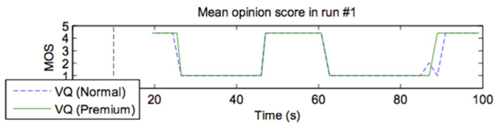
MOS score of the RTP video in both runs

Figure 1 - Application differentiation test

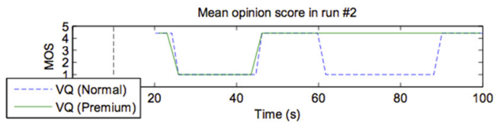
Run #1 was performed without the application-based priority handling, whereas run #2 was performed with it. The dashed vertical lines indicate the start of the video and bulk streams, respectively at 10s and 60s. It is clear from Figure 1 that with the traffic control off, the video quality quickly turns unacceptable, while when it is on, the quality remains at acceptable levels.

The second set of tests involved two video streams belonging to different subscriber classes. The link bandwidth was set so that one flow could be served without problems, but two flows would congest it. **Figure** 2 shows the results obtained. In the first run, the quality of both streams suffers, as expected, since the link cannot support both at their peak rates (Figure 2a). **Figure** 2b shows the effect of the subscriber class differentiation at work, and it is easy to see that the premium user enjoys a significantly better quality than the normal user. The reader may notice that there is a period in which the premium user will also suffer from a lower quality in a first instance in run #2. This is due to a trade-off between the size of the time window over which the MOS is estimated, and the estimation's accuracy. In practical use, a smaller window would probably be useful to avoid the user stopping the streaming due to the lowered quality.

This leads us to the thirds test set, regarding the reaction times of the system. The fastest performance achieved resulted in flows being promoted to a higher-priority class in 2.8s on average over 40 test runs (recall that all flows start out in the same FIFO queue by default if there's no contention). The total reaction time was of 4.0s. As mentioned before, however, these smaller values impose a trade-off in the QoS calculations, which become noisier as a consequence of having fewer samples.



(a) MOS score of the RTP videos in the first run



(b) MOS score of the RTP videos in the second run

Figure 2 - Subscriber differentiation test

The final set of tests was related to the admission control. The QoE of the flows in the system was averaged over a 30s sliding window, in order to avoid noise in the measurements. In this setup, two streams were started at different times, and the link bandwidth was set low, so that the quality of the first stream was acceptable, but not good.

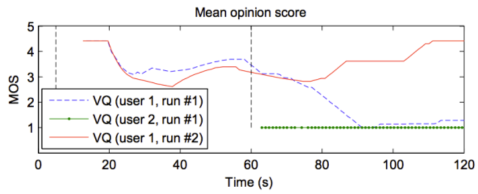


Figure 3 – Admission control test

In Figure 3, we can see that in run #1, when the admission control is not enabled, the start of the second stream results in a completely unacceptable quality for both streams. Note that actual degradation is sharper than it appears in the plot, as the plot is smoothed by the 30s averaging window. In run #2, with admission control enabled, when the second flow starts it is immediately dropped, as the quality of the first flow is below the activation threshold. Thus, the user watching that stream attains an acceptable quality throughout the whole period.

4. Conclusions

We have proposed a multi-faceted approach to QoE-based traffic control by considering different subscriber and application types and using them to perform admission control and traffic differentiation. The results obtained show a clear QoE improvement for OTT video streams when the proposed mechanisms are in place instead of a simple best-effort policy. Further work on this subject includes the extension and refinement of the traffic classification mechanism used to work on adaptive HTTP-based video streaming schemes, as well as the development of suitable parametric QoE models for them.

References

[1] Cisco Systems, Inc., "Cisco Visual Networking Index: Forecast and Methodology, 2011-2016," May 2012.

[2] Janne Seppänen and Martín Varela, "QoE-Driven Network Management for Real-Time Over-the-Top Multimedia Services," in *IEEE WCNC*, Shanghai, 2013, Accepted for publication.

[3] D. Waring, G. Lapiotis, J. Lyles, and R. Vaidyanathan K. Kerpez, "IPTV Service Assurance," *IEEE Communications Magazine*, vol. 44, no. 9, pp. 166-172, September 2006.

[4] M. Garcia, M. Atenas, and A. Canovas J. Lloret, "A QoE management system to improve the IPTV network," *International Journal of Communication Systems*, vol. 24, no. 1, pp. 118-138, 2011.

[5] K. Haugene and A. Jacobsen, "Network based QoE optimization for "over the top" services," *Norwegian University of Science and Technology*, Department of Telematics, 2011.

[6] A. Ksentini, C. Viho, and J.-M. Bonnin K. Piamrat, "QoE-aware admission control for multimedia applications in IEEE 802.11 wireless networks," in *IEEE VTS - Vehicular Technology Conference (Fall)*, 2008, pp. 1-5.

[7] M. Hirvonen, "Two-phased network traffic classification method for quality of service management," University of Oulu, Dpt. of Electrical and Information Engineering, 2009.

[8] P. E. McKenney, "Stochastic Fairness Queueing," in *IEEE INFOCOM'90*, vol. 2, 1990, pp. 733-740.

[9] Martín Varela, "Pseudo-subjective Quality Assessment of Multimedia Streams and its Applications in Control," *Université de Rennes 1*, Doctoral Thesis Nov. 2005.

[10] M. Devera. (2003) HTB Home. [Online]. http://luxik.cdi.cz/~devik/qos/htb/

[11] B. Hubert. Linux Advanced Routing & Traffic Control HOWTO. [Online]. http://lartc.org/lartc.html#LARTC.SFQ

Dr. Martín Varela received his PhD and MSc from the University of Rennes 1 (Rennes, France), in 2005 and 2002 respectively. He has been an ERCIM fellow, and spent time at SICS and VTT, where he has been a Senior Scientist since 2007. His research interests lie in the QoE domain, both in QoE for real-time multimedia services and for generic services.

Janne Seppänen received his Bachelor's Degree in University of Oulu in 2010 and his Master's Degree two years later. In his studies, he specialized in computer engineering and signal processing. He currently works at Technical Research Centre of Finland (VTT) as a research scientist, covering topics such as Quality of Experience, network management, traffic classification and network traffic measurement.