Objective and Subjective Evaluation of QOAS Video Streaming over Broadband Networks

Gabriel-Miro Muntean, Member, IEEE, Philip Perry, Member, IEEE Liam Murphy, Member, IEEE

Abstract—This paper presents objective and subjective testing results that assess the performance of the Quality-Oriented Adaptation Scheme (QOAS) when used for high quality multimedia streaming over local broadband IP networks. Results of objective tests using a QOAS simulation model show very efficient adaptation in terms of end-user perceived quality, loss rate, and bandwidth utilization, compared to existing adaptive streaming schemes such as LDA+, and TFRCP. Subjective tests confirm these results by showing high end-user perceived quality of the QOAS under various network conditions.

Index Terms—Adaptive multimedia streaming, feedback control, grading scheme, end-user quality.

I. INTRODUCTION

Lately there is a significant increase in the need for delivering multimedia-based services to home residences and business premises [1]. The pressure put on the delivery networks by the consequent rise in the volume of data carried has further increased by the choice for high quality content that determined the traffic to build up even more. Bursty losses, or excessive and extremely variable delays, caused by this increased traffic have a devastating effect on multimedia delivery over IP networks by severely affecting the end-users' perceived quality. Regardless of the infrastructure architecture used for delivering rich content multimedia-based services [2], the service providers and network operators aim at increasing its utilization and thus their revenues. On the other hand the customers always want the best quality for the services at the lowest price possible.

The end-to-end application-level adaptive control solution the Quality-Oriented Adaptation Scheme (QOAS) - that was proposed in [3], [4] and described in [5], [6], [7], balances these opposing requirements and works best in increased traffic conditions. The adaptation is based on a client-located grading scheme that maps some network-related parameters' values and variations to application-level scores that describe the quality of delivery. In order to maximize the quality of service in existing conditions, estimates of end-user perceived quality are actively considered during this grading process. The computed quality scores are used by a server-side feedback-controlled QOAS mechanism to take adaptive decisions.

Results of extensive testing that assess QOAS when delivering multimedia streams over local broadband IPnetworks [1] are presented and discussed here, extending results published previously in [8]. They illustrate QOAS performance and its potential benefits for delivering multimedia-based services to the customers. These tests involve both simulations and subjective perceptual testing and their results are presented in sections IV and V. Section II discusses briefly some previous works that were proposed in order to provide certain level of quality of service (QoS) when delivering multimedia streams and gives details about adaptive solutions. It also presents some approaches for assessing the end-user perceived quality involving objective metrics and subjective tests. Section III gives some details about QOAS and its principle before the testing results are presented in detail and commented in the following sections. At the end of the paper, performance analysis, conclusions and future work directions are presented.

II. RELATED WORK

Extensive research was focused on devising solutions for providing certain level of QoS when delivering multimedia data over best-effort networks that would also take into account variable delivery conditions. Various technologies and architectures and different approaches were proposed and among the best known are those based on bandwidth overprovisioning, traffic engineering, QoS architectures and adaptive solutions.

Bandwidth over-provisioning [9] involves allocating statically more bandwidth than the expected network peak requirements. It increases the probability of avoiding congestion-related problems, but provides no guarantees, especially during peak-hours and with very bursty traffic, while wasting resources most of the time.

Traffic engineering is concerned with planning, design, monitoring and management of networks and their traffic in order to allow for the most efficient transport possible. Although traffic engineering-based solutions (e.g. Integrated Services (IntServ) - RFC 1633 [10], Differentiated Services (DiffServ) - RFC 2475 [11], Multiprotocol Label Switching (MPLS) - RFC 3031 [12], etc.) have unchallenged merits, there are significant concerns regarding their complexity, deployment costs and some issues related to security, size of targeted networks and capability of reaction in really congested conditions.

Manuscript received September 17, 2004 and revised July 1, 2005. It extends a paper published in the Proceedings of the IEEE/IFIP International Conference on Management of Multimedia Networks and Services, San Diego, CA, USA, October 2004. The Associate Editor who handled this paper was Juergen Schoenwaelder.

G.-M. Muntean is with the School of Electronic Engineering, Dublin City University, Glasnevin, Dublin 9, Ireland (e-mail: munteang@eeng.dcu.ie)

P. Perry is an independent consultant radio engineer (philip@perryradio.com) L. Murphy is with the Department of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland (liam.murphy@ucd.ie)

QoS architectures provide a unifying framework for different aspects concerned on QoS provisioning, ranging from user requirements through operating system and hardware characteristics to network capability and performance [13], [14]. Proposals include the Lancaster QoS-Architecture (QoS-A) [13], [14], OSI QoS Framework Model [15], [16], Tenet Approach [17], Heidelberg HeiProjects [18], etc.) which can provide good results, but are complex and relative expensive to deploy.

The **adaptive schemes** for multimedia streaming take the distribution networks as they are and provide the least complex and the most flexible mechanisms for trying to provide certain level of QoS in existing network conditions. They adjust the bandwidth used by the adaptive applications according to the existing network conditions, increasing or decreasing the transmission and/or encoding rates. The design alternatives explored differ on how some important issues are taken into consideration. Some of these issues are:

- Signaling or feedback mechanism used to inform the sender and/or receiver about the current network conditions,
- Specific adaptive mechanisms used in response to this information,
- Localization of this adjustment mechanism,
- Responsiveness of the congestion control scheme in detecting and reacting to network conditions,
- Capability of the scheme to accommodate heterogeneous receivers that may differ in their connectivity to the network, the amount of traffic to their delivery paths, their need for quality
- Scalability of the control mechanism to a high number of receivers
- Sharing of bandwidth with competing traffic of different type (particularly with TCP)
- Perceived quality of adapted multimedia streams.

A. Adaptive Streaming Solutions

Extensive research has focused on proposing different adaptive schemes based on rate control and various directions have been taken. They were mainly classified in the literature [19], [20], [21] according to the place where the adaptive decision is taken.

Source-based adaptive control techniques require the sender to respond to variations in the delivery conditions. Among them there are solutions based on *probing tests* that try to estimate the available bandwidth while maintaining the loss rate below a certain threshold [22], [23].

Other solutions follow a *throughput model* that determines the transmission rate in certain conditions. The TCP-Friendly Rate Control Protocol (TFRCP)-based adaptive scheme [24] relies on the TCP model proposed in [25], the Time-based Model TCP-Friendly Rate Control (TMRC) [26] on the TCP rate estimation model proposed in [27], whereas the Loss-Delay Adjustment Algorithm (LDA) [28] uses an original model for rate adaptation.

A third direction that *relies on heuristic knowledge*, *experimental testing* and *models* encompasses many of the proposed schemes. Among the most significant are the LossDelay-based Adaptation Algorithm (LDA+) [29] that extends LDA; the Rate Adaptation Protocol (RAP) [30] which uses a similar approach to TCP's AIMD adaptation; Layered Quality Adaptation (LQA) [31] that bases its rate control on a layered approach; and the scheme described in [32] that bases its adaptation on information about the network state acquired by a TCP-like mechanism.

Receiver-based schemes provide mechanisms that allow for the receivers to select the service quality and/or rate, such as Receiver-driven Layered Multicast (RLM) [33] and Receiverdriven Layered Congestion Control (RLC) [34]. Among the **hybrid adaptive mechanisms** that involve both the sender and the receiver in the adaptation process, the TCP Emulation At Receivers (TEAR) scheme was described in detail in [35]. **Transcoder-based solutions** focus on matching the available bandwidth of heterogeneous receivers through transcoding or filtering, and significant solutions were presented in [36], [37].

Commercial adaptive streaming solutions like Real Networks' SureStream [38] and Microsoft's Multimedia Multi-bitrate (MBR) solution [39] are proprietary and detailed technical information has never been revealed. However the available information states that they were specially designed to allow for adaptations at very low bit-rates, unlike QOAS which addresses high quality high bit-rate video streaming.

B. End-user Perceived Quality Evaluation

Significant research was performed in the area of end-user perceived quality assessment and two main directions were explored: objective and subjective testing.

Objective methods aim at determining the quality of a multimedia sequence in the absence of the human viewer. They are classified in [40] in *out-of service methods* (used offline when the original sequence is fully available and no time constraints are imposed) and *in-service methods* (used during streaming when the original clip is not available and have strict timing requirements).

A classification based on the existence of the original multimedia stream [41] distinguishes three approaches: *full reference methods* (based on picture comparison), *reduced reference solutions* (relies on feature extraction) and *no reference methods* (also called single-ended). Only the last category of methods is useful for in-service applications. Most of these methods have associated metrics in order to allow for quantification of the end-user perceived quality.

Other works such as [42, 43] divide the metrics associated to these objective methods into **mathematical-based** (rely on mathematical formulae or on functions based on intensive psycho-visual experiments) and **model-based** (based on complex models of the human visual system).

The full-reference Peak Signal to Noise Ratio (PSNR) [44] and Weighted Signal to Noise Ratio (WSNR) [44] and the noreference Picture Appraisal Rating (PAR) [45] proposed by Snell & Wilcox¹ for MPEG-2 videos are among the mathematical metrics. Although they seem appropriate and are very simple, many studies [40, 46] have shown that PSNR and WSNR are poorly correlated to human vision.

¹ Snell & Wilcox, Web Site, http://www.snellwilcox.com

Among the model-based metrics are PQR [47], KDD [48], DVO [49], VOM [50], PVOM [51], PDM [52] and MPOM [53]. The Picture Ouality Rating (POR) [47] is a full reference metric based on Tektronix²/Sarnoff³ Human Vision Model that relies on the proprietary JNDmetrix (Just Noticeable Difference)⁴. Kokusai Denshin Denwa (KDD) Research and Development Laboratories⁵ proposed a proprietary fullreference model [48] based on mean square error that is weighted by a set of sequential Human Visual Filters applied at pixel, block, frame and sequence levels. The Digital Video Quality (DVQ) metric [49] and the Video Quality Model (VQM) [50], full-reference metrics proposed by NASA⁶ and the Institute for Telecommunication Sciences, NTIA⁷ USA respectively, are subject to U.S. patents [54], 55] and unfortunately their usage involves licensing costs. The Perceptual Video Quality Measure (PVQM) [51] uses the same approach for measuring video quality as used for speech in the Perceptual Speech Quality Measure (PSQM), standardised by the ITU-T [56]. The Perceptual Distortion Metric (PDM) [52], proposed by L'Ecole Polytechnique Federale de Lausanne (EPFL) Switzerland, is based on a spatio-temporal model of the human visual system. Although complex, this full-reference metric's greatest advantage is that details about it are made public. Researchers from EPFL have also proposed the Moving Pictures Quality Metric (MPQM) [53] which is a full reference video quality metric based on an multi-channel human visual model that takes into consideration contrast sensitivity and intra-channel masking and the no-reference MPQM (Q) [46] that describes the joint impact of MPEG rate and data loss on video quality. The latter is highly useful for in-service applications and is also used by QOAS [5].

ITU-T Video Quality Expert Group⁸ has studied extensively objective metrics proposals for standardization and concluded that no metrics outperforms the others in all conditions. In consequence currently no objective solution is able to fully replace subjective testing [44] which are necessary.

Subjective tests as defined by ITU-R BT.500 [57] have been used for many years in order to assess the quality of television pictures. In the area of telecommunications five major ITU-T recommendations concern subjective testing: P.910 [58] - one way video test methods, P.911 [59] - quality assessment methods for multimedia applications, P.800 [60] – conditions for audio content testing, P.920 [61] - conversation quality assessment and P.930 [61] - video impairment reference system. Among them the first two mostly present recommendations about methods, systems, clip contents and environment conditions for subjective testing and scales for assessing the end-user perceived quality while viewing multimedia clips.



Fig. 1 QOAS adaptation principle, illustrated for QOAS-based pre-recorded multimedia streaming

III. QUALITY-ORIENTED ADAPTATION SCHEME (QOAS)

Although the adaptive schemes presented in the "Related Work" section have shown good adaptation results in certain scenarios, their adjustment policies are not directly related to the quality of the streaming process as perceived by the viewers. Unlike them, QOAS bases its adaptation process on estimates of the end-user perceived quality made at the receiver. This perceived quality is estimated in-service using the no-reference moving picture quality metric-Q proposed in [46] that describes the joint impact of MPEG rate and data loss on video quality. More details about Q and how QOAS makes use of it are presented in [5].

QOAS is distributed and consists of server-side and clientside components. It makes use of a client-located Quality of Delivery Grading Scheme (QoDGS) and of a Server Arbitration Scheme (SAS) that co-operate in order to implement the feedback-controlled adaptation mechanism. The QOAS principle is schematically illustrated in Figure 1 for pre-recorded multimedia streaming used for Video-on-Demand (VoD) services, and is briefly described next.

A. Principle of Quality-Oriented Adaptive Scheme

Multimedia data is received at the client where the QoDGS continuously monitors both some network-related parameters such as loss rate, delay and jitter and the estimated end-user perceived quality. According to their values and variations, OoDGS grades the quality of delivery (OoD) in terms of application-level quality scores (QoD_{Scores}) that are sent to the server as feedback. These scores are analyzed by the SAS that may suggest taking adaptive decisions in order to maximize the end-user perceived quality in existing delivery conditions. These decisions affect an internal state defined for the QOAS server component that was associated with the streamed multimedia clip's quality as shown in Figure 1. The figure presents the five-state quality model used during testing with the following states: excellent, good, average, poor and bad. Between adjacent states the adaptation step is 0.5 Mbps in the experiments described in this paper. Any OOAS server state modification affects the multimedia data transmission rate. For example, when increased traffic in the network affects the client-reported quality of delivery, SAS switches to a lower quality state. This results in a reduction in the quantity of data sent, thus helping to improve the situation. This is performed

² Tektronix, http://www.tek.com

³ Sarnoff, http://www.sarnoff.com

⁴ Just Noticeable Difference Metrics, http://www.JNDmetrix.com

⁵ Kokusai Denshin Denwa Research and Development Laboratories, KDDI Corporation, http://www.kddilabs.jp/english

⁶ The National Aeronautics and Space Administration (NASA), USA, http://www.nasa.gov

⁷ Institute for Telecommunication Sciences, National Telecommunications and Information Administration (NTIA), USA, http://www.its.bldrdoc.gov

⁸ Video Quality Experts Group (VQEG), http://www.its.bldrdoc.gov/vqeg

because research has shown [63] that viewers prefer a controlled reduction in multimedia quality to the effect of random losses on the streamed multimedia data. In improved delivery conditions, the QOAS server component gradually increases the quality of the transmitted stream and therefore the transmission rate. In the absence of loss this causes an increase in end-user perceived quality.

B. Quality of Delivery Grading Scheme (QoDGS)

QoDGS maps some transmission related parameters values and variations and estimates of end-user perceived quality into application-level scores that describe the quality of delivery. It monitors some parameters such as delay, jitter and loss rate, computes estimates of end-user perceived quality using Q and analyses their short-term and long-term variations. Short-term monitoring is important for learning quickly about transient effects, such as sudden traffic changes, and for quickly reacting to them. The long-term variations are monitored in order to track slow changes in the overall delivery environment, such as new users in the system. These shortterm and long-term periods are set to be an order and two orders of magnitude (respectively) greater than the feedbackreporting interval in the experiments described here.

In the first of QoDGS's three stages, instantaneous values of the monitored parameters are saved in different length sliding windows and their short-term and long-term variations are assessed. At the same time, session-specific lower and higher limits are maintained for each parameter, allowing for corresponding partial scores to be computed in comparison with them. In the second stage, the relative importance of all the monitored parameters in this delivery infrastructure is considered (by weighting their contributions) and the partial scores are used to compute short-term (QoD_{ST}) and long-term (QoD_{LT}) quality of delivery grades. This second stage also takes into account estimates for short-term and long-term enduser perceived quality. In the third stage, QoD_{ST} and QoD_{LT} are weighted to account for their relative importance and the overall client score (QoD_{Score}) is computed.

Extensive tests were performed in order to make sure that the design of QODGS ensures that best results will be obtained in terms of adaptiveness, responsiveness to traffic variations, stability, link utilization, and end-user perceived quality in local broadband IP-networks. A detailed presentation of QoDGS is given in [5].

C. Server Arbitration Scheme (SAS)

SAS takes adaptive decisions based on the values of a number of recent feedback reports, in order to minimise the effect of noise in the QoD_{Scores} . This arbitration process is *asymmetric*, requiring fewer feedback reports to trigger a decrease in quality than for a quality increase. This ensures a fast reaction during bad delivery conditions, helping to eliminate their cause and allowing the network conditions to improve before any quality upgrade. These adaptive decisions are taken to maintain system stability by minimising the number of quality variations. The late arrival of a number of feedback messages is considered as an indication of network congestion, and triggers quality degradations. This permits the streaming scheme to work even if feedback is not available. More details about SAS are presented in [5].



Fig. 2. "Dumbbell" topology for NS-2 simulation tests

IV. OBJECTIVE TESTING RESULTS

In order to test QOAS performance when delivering multimedia clips in local multi-service broadband IP-networks to home residences and business premises, QOAS was implemented by both a simulation model, built using Network Simulator 2 (NS-2) [64], and a prototype system, built using Microsoft Visual C++ 6.0. The simulation model was used for objective testing whereas the prototype system was used for subjective assessment of the end-users' perceived quality. This section is focused on presenting objective testing results.

The objective testing employs NS-2 simulations in order to assess the QOAS performance. The simulation setup requires a network topology, simulation models, multimedia clips, simulation scenarios and performance assessment principles. These issues and the simulation results are presented next.

A. Network Topology

The NS-2 simulations use a "Dumbbell" topology that assumes a single shared bottleneck link with characteristics as in Figure 2. The 100 ms latency was chosen so that the adaptation of the feedback-based schemes in highly loaded delivery conditions was tested. The sources of traffic, including QOAS server application instances and a source of multimedia-like background traffic are located on one side of the bottleneck link, and the receivers are on the other side. The links are provisioned such as the only significant delays and packet drops are caused by congestion that occurs on the bottleneck.

B. Simulation Models

For testing QOAS a simulation model that implements the mechanism described in section III was built, using a fivequality state model for the server. The SAS upgrade period was 6 s and the downgrade one was 1 s. The QoDGS shortterm and long-term periods were set to 1 s and 10 s, respectively.

When comparing QOAS to other adaptive schemes, NS-2 models for TFRCP [24] and LDA+ [29] were used and maximum rate of 4 Mb/s was imposed for consistency.

TFRCP uses estimates of round-trip delay and loss rates to determine the adaptive policy. When loss occurs, the rate of transmission is limited to the one computed by the TCP model [25]. In case of no loss, the current rate is doubled. This TFRCP model uses 5 s for rate update intervals, as suggested in [24] for latencies greater than 0.1s, as in this setup.

TABLE I
PEAK/MEAN BIT-RATE RATIO FOR ALL QUALITY VERSIONS OF THE
MULTIMEDIA CLIPS USED DURING SIMULATIONS
· · · · ·

Quality Version (average rate) Clip Name	2.0 Mb/s version	2.5 Mb/s version	3.0 Mb/s version	3.5 Mb/s version	4.0 Mb/s version
diehard1	7.48	7.43	6.31	5.65	4.06
roadtoeldorado	6.91	6.51	6.23	6.12	6.05
dontsayaword	5.56	4.51	4.36	4.08	3.56
jurassic3	4.83	4.38	4.04	3.71	3.41
familyman	3.99	3.67	3.42	3.09	2.93

LDA+ is an AIMD algorithm based on estimates of network condition and bandwidth share used. In zero loss periods, the sender increases its rate with minimum from an estimated bandwidth share rate increase, a bottleneck bandwidth share rate limit, and a corresponding TCP rate update. In nonzero loss periods, the server reduces its rate by a value that depends on the current rate and the rate determined by the TCP model [25]. The LDA+ implementation used an RTCP feedback interval of 5 s as suggested in [29].

C. Multimedia Clips

Five video sequences were selected from movies with various types and representing different classes of clips in terms of the degree of motion content: diehard1 - high motion content, jurassic3 and dontsayaword - average, familyman low, whereas *roadtoeldorado* is a typical cartoon sequence. The clips were MPEG-2 encoded at five rates between 2 Mb/s and 4 Mb/s using the same frame rate (25 frames/sec) and the same IBBP frame pattern (9 frames/GOP). The resolution was constant for all sequences: 320 x 240. Traces were collected, associated with QOAS server states and used during simulations. Statistics related to the ratio between the peak and mean rates for each version of the multimedia sequences used during simulations are presented in Table I. Peak/mean bit-rate ratios are closely related to both the motion complexity and type of multimedia sequences and are the cause for the burstiness of transmissions.

D. Simulation Scenarios and Results

Simulations involved streaming each multimedia clip indicated in Table I for 500 s, but 50 s long transitory periods at the beginning and the end were not considered when analysing the results. Since multimedia is expected to account



Fig. 3. Multimedia-like background traffic on top of 95.5 Mb/s CBR traffic

TABLE II PERFORMANCE COMPARISON WHEN STREAMING *diehard1* WITH OOAS, TFRCP AND LDA+ RESPECTIVELY

Avg. Tx. Rate (Mb/s)	Avg. Loss (%)	Avg. Quality (1-5)	Avg. Utilis. (%)				
3.21	0.013	4.42	99.91				
3.16	1.057	3.79	99.88				
2.95	1.465	3.77	99.67				
	Avg. Tx. Rate (Mb/s) 3.21 3.16 2.95	Avg. Tx. Rate (Mb/s) Avg. Loss (%) 3.21 0.013 3.16 1.057 2.95 1.465	Avg. Tx. Rate (Mb/s) Avg. Loss (%) Avg. Quality (1-5) 3.21 0.013 4.42 3.16 1.057 3.79 2.95 1.465 3.77				

for the majority of traffic in local multi-service broadband IP networks, the complex multimedia-like background traffic presented in Figure 3 is used. This traffic simulates possible user interactions such as consecutive *play* commands that increase the traffic in a staircase-up manner, different frequency *pause-play* interactions applied on different rate clips and consecutive *stop's*. In order to create highly loaded network conditions, CBR-UDP background traffic with a rate of 95.5 Mb/s is generated using NS-2. This traffic represents



Fig. 4a. Average bitrate variation when streaming *diehard1* sequence with QOAS triggered by the background traffic



Fig. 4b. Average bitrate variation when streaming *diehard1* sequence with TFRCP triggered by the background traffic



Fig. 4c. Average bitrate variation when streaming *diehard1* sequence with LDA+ triggered by the background traffic



Fig. 5c. Loss rate variation experienced when streaming *diehard1* sequence with LDA+

Time (s)

the well-multiplexed aggregation of a high number of data

flows of different types commonly expected in IP networks. Table II presents comparative performance statistics gathered when streaming *diehard1* using QOAS, TFRCP and LDA+, respectively in these traffic conditions. The performance was assessed in terms of average bit-rate, enduser perceived quality, loss rate and infrastructure utilization. End-user quality is computed using the no-reference metric Q [46] and is expressed on the ITU-T five-point scale [58].

Since QOAS maintains very low loss rate (0.01 %) by successfully adapting even to most difficult background traffic variations, the consequent average end-user perceived quality is between "good" and "excellent" quality level (4.42). Higher loss rates than 1% are experienced when using both TFRCP and LDA+, determining decreases of the end-user perceived quality much below the "good" perceptual level.

Tests were performed using multimedia clips with different motion content (see Table I) and the results were similar. As an example, this paper presents a comparison between performance-related results obtained when streaming a single multimedia sequence with very complex motion content *diehard1* using QOAS, TFRCP and LDA+. The performance



Fig. 6a. End-user perceived quality when streaming *diehard1* sequence using QOAS



Fig. 6b. End-user perceived quality when streaming *diehard1* sequence using TFRCP



Fig. 6c. End-user perceived quality when streaming *diehard1* sequence using LDA+

is assessed in terms of rate adaptation to background traffic variation (Figures 4a, 4b and 4c), loss rate (Figures 5a, 5b and 5c), end-user perceived quality (Figures 6a, 6b and 6c), and link utilization (Figures 7a, 7b and 7c). Next these results are presented and commented in details.

QOAS successfully adapts to the staircase-up increase in the background traffic that exceeds the available bandwidth devised from 0 s to 150 s, reducing the quantity of data transmitted and completely avoiding losses. However, as Figures 5b and 5c show, during both TFRCP and LDA+-based streaming, losses occur before the rate is reduced causing significant degradations of end-user perceived quality.

When the background traffic varies in a periodic manner with steps comparable to the adaptation step of 0.5 Mb/s (see Table I), QOAS obtains better results in terms of perceived quality in comparison to both other solutions due to its conservative policy of slowly increasing the transmission rate to a level determined according to long-term information it maintains. Both LDA+ and TFRCP use a more aggressive manner of recovery after network problems and increase their transmission rate faster as plotted in Figures 4b and 4c. This policy achieves in generally high throughput, but when the background traffic varies sharply like in this situation, it leads to packet loss. These lossy periods can be observed in Figures 5b and 5c.

The effect of a steep increase in the background traffic when the system is already heavily loaded is tested at 250 s and 360 s. QOAS performs significantly better that both TFRCP and



Fig. 7a. Link utilization when streaming *diehard1* sequence using QOAS



Fig. 7b. Link utilization when streaming *diehard1* sequence using TFRCP



Fig. 7c. Link utilization when streaming *diehard1* sequence using LDA+

LDA+-based adaptations, reacting much faster to the sharp change in traffic. This minimizes the losses and therefore reduces the period when the perceived quality is degraded from 20 s in TFRCP case (see Figures 5b and 6b) and 17 s in LDA+ case (see Figures 5c and 6c) to only 1.2 s (see Figures 5a and 6a).

At the end of this set of tests, the effect stopping the multimedia cross traffic has on multimedia streaming with different approaches was tested. All the adaptive schemes increased their rates to compensate for the decrease in background traffic, but TFRCP and LDA+ did this faster than QOAS, as it could be notices from Figures 4a, 4b and 4c. However, the difference in the perceived quality between the consequent results was less than 2%. A comparison between instantaneous estimations of end-user perceived quality using Q metric expressed on the ITU-T P.910 five-point scale in all the three cased studied is presented in Figures 6a, 6b and 6c.

As similar results were obtained when the other multimedia clips with different degree of motion content were used for streaming, it can be concluded that the QOAS-based solution showed superior performance to both TFRCP and LDA+. QOAS reacts quickly to changes in network traffic, reducing the quantity of the transmitted data, both preventing and minimizing losses, if they occur. Therefore the consequent

 TABLE III

 AVERAGE BITRATE AND LOSS RATE EXPERIENCED

 DURING SUBJECTIVE TESTS Test1 AND Test2

Sequence	Test 1		Test 2	
	Avg. Rate (Mbps)	Avg. Loss Rate (%)	Avg. Rate (Mbps)	Avg. Loss Rate (%)
diehard1	3.27	0.26	2.89	0.33
dontsayaword	3.25	0.21	2.88	0.32
familyman	3.29	0.27	2.91	0.29
roadtoeldorado	3.26	0.23	2.87	0.28



Fig. 8. Testbed setup for subjective testing

end-user perceived quality was much higher than in the other cases when it even reached the "very annoying" level for long periods. QOAS's more conservative upgrade approach pays off if unexpected delivery problems occur.

In terms of link utilization, as it could be seen from the results presented in Table II, all the solutions have highly performed, although on average QOAS slightly out-performs the other schemes. Looking at instantaneous values for link utilization, with QOAS the variation is smoother than with the other schemes that vary their rate more during the session, sometimes achieving sub-optimal utilization and other times recording loss.

V. SUBJECTIVE TESTING RESULTS

Subjective tests were performed in order to verify the objective end-user quality results obtained during simulations. They have involved the prototype system and 60 s long multimedia sequences taken from the same movies with different motion content used during simulations (see Table I). Increased traffic conditions were emulated using the NistNet network emulator [65] determining QOAS-based adaptations and consequent variations in the viewers' perceived quality.

The testbed presented in Figure 8 was set up such as it involves a QOAS server application deployed at the Local Server and a QOAS client application run on the Local Client machine. The client application makes use of a Canopus⁹ Amber MPEG decoder card and the corresponding SDK. No error control and error concealment methods were employed. Testing conditions suggested in the ITU-T R. P.910 [58] and ITU-T R. P.911 [59] were ensured and the Single Stimulus Method with explicit reference was selected as testing methodology for two perceptual tests. These aimed at testing

⁹ Canopus United Kingdom, http://www.canopus-uk.com

TABLE IV SUBJECTIVE TEST RESULTS: MEAN PERCEIVED QUALITY SCORES FOR Test1 AND Test2 Sequence Motion Content/ Type Test 1 Test 2 4.00 4.22 diehard1 High / Movie dontsayaword Average / Movie 4.18 3.98 familyman Low / Movie 4.21 4.24 roadtoeldorado Average / Cartoons 3.74 3.85

the subjects' perceived quality when using QOAS for streaming in very difficult delivery conditions, as shown by the simulations (see section IV).

The effects of consecutive *play* commands in the delivery system that are emulated by background traffic that varies in a staircase-up manner are tested in *Test1*. The effects of periodic variation of traffic with steps of 0.7 Mb/s, higher than the adaptation step of 0.5 Mb/s, are assessed in *Test2*. Figure 9 and Figure 10 show both the background traffic variations and the consequent QOAS rate adaptations during these tests, when the *diehard1* clip was selected for streaming. Similar results were obtained when the other clips were used. Table III presents average bitrate of the streamed multimedia clip and average loss rate recorded during the streaming session.

In each of the two tests 42 subjects, aged between 18 and 48, graded the quality of each streamed clip on the 1-5 ITU-T R. P.910 scale [58]. No fractional grades were accepted. Among the subjects, 19 and 16 in the first and the second tests respectively wore glasses or contact lenses and none had other visual impairments that may affect their perception of multimedia quality. From the subjects, 23 and 21 respectively were familiar with multimedia streaming, 1 and 2 respectively have considered themselves experts.



Fig. 9. Test 1: QOAS bit-rate adaptation with background traffic variation when streaming *diehard1*



Fig. 10. Test 2: QOAS bit-rate adaptation with background traffic variation when streaming *diehard1*

The results presented in Table IV show how QOAS streaming was very appreciated by the test subjects, scoring on average above 4, the "good" quality level on the ITU-T scale, for all the movies and close to the "good" level for the cartoons sequence.

The results of *Test1* suggest that the higher the motion complexity of a sequence, the more the end-user perceived quality was affected by the effect of the difficult delivery conditions. The fact that the users' subjective appreciation in loaded delivery conditions was lower is confirmed by an ANOVA test, which indicated that the results are significantly different (p < 0.05). However, during *Test2* when the delivery conditions have triggered loss, the viewers' perceived quality was affected independent from the motion content as shown in Table IV. This finding was supported by an ANOVA test that found the results significantly different (p < 0.05).

Although the results of the second set of subjective test seem higher than those of the first set of tests, by performing t-tests on *Test1* and *Test2* results for each multimedia sequence involved in testing, the null hypothesis that there is no statistical difference between the results of *Test1* and *Test2* respectively cannot be rejected. This finding is stated with a very high level of confidence of 99% (significance level $\alpha =$ 0.01).

At the same time there is a significant statistical difference between the subjective scores obtained for the clips that contain movie scenes and the cartoons clip. This result was confirmed by paired t-tests that were performed for each movie sequence and the cartoons sequence with a significance level of $\alpha = 0.01$. A potential cause might be the different MPEG-2 encoding output for the cartoons sequences as shown in Table I. Unlike for the movie content, for cartoons content the peak/mean ratio computed in relation to the size of the encoded frames does not significantly increase with the decrease in the average encoding bit-rate. Also the content with many colors and clearly defined edges might be more affected in terms of the end-user subjective quality corrupted during streaming.

In conclusion, although slightly lower than the simulation test results obtained in the same conditions (for example when streaming the *diehard1* sequence the mean scores were 4.42 and 4.22 respectively) the subjective test results verify them and confirm the very good performance of QOAS.

VI. PERFORMANCE ANALYSIS

The significant advantages of a QOAS-based solution come with a cost in terms of extra processing requirements and some bandwidth used for feedback.

The fact that this processing is distributed among the QOAS clients whose QoDGSs monitor and grade the quality of streaming at the receivers, significantly reduces the load of the QOAS server machine that runs only the SAS. The QOAS server has only to acquire the client transmitted QoD_{Scores} , to process them (this can be performed incrementally) and to take adaptive decisions (this does not involve excessive CPU load).

Regarding the feedback, it is significant to mention that each feedback report consists only of a QoD_{Score} . If RTCP packets

are used, for standard values for the headers' sizes (20 Bytes – IP header, 8 Bytes – UDP header, 8 Bytes – RTCP receiver report packet header) and for a 4-Byte payload, the feedback packet size becomes 40 Bytes long. For a very low interfeedback transmission time of 0.1 sec the bandwidth used by feedback for a single client becomes $BW_{feedback} = 400$ Bytes/s. Since QOAS was designed for local broadband multi-service IP-networks, this represents an insignificant bandwidth usage. For example over 300 customers can be served simultaneously via a gigabit Ethernet infrastructure and will consume only 0.1 % of the available bandwidth for feedback. This value is significantly lower than the upper limit of 5 % of bandwidth suggested by RTP/RTCP in [66].

VII. CONCLUSIONS AND FURTHER WORK

The Quality-Oriented Adaptation Scheme (QOAS) is an end-to-end application-level solution for streaming multimedia that considers the end-user perceived quality as an active factor in the adaptation process. The scheme is tested in conditions expected for delivering multimedia-based services to residential homes or business premises via a local broadband multi-service IP network.

Simulation-based objective tests have shown very good performance of QOAS, assessed in terms of remote user perceived quality, average loss rate and network infrastructure utilization when streaming multimedia in loaded network conditions and in the presence of highly variable multimedia-like background traffic. The perceived quality was between the "good" and "excellent" ITU-T quality levels, the loss rate was around 0.01 % and the utilization greater than 99.9 %. These results show better performance than those obtained when other adaptive schemes such as TFRCP and LDA+ were tested in the same conditions. Subjective tests performed in difficult emulated traffic conditions verify these results.

These results highly recommend QOAS as a very efficient solution for delivering good quality multimedia-based services to customers in local broadband IP-networks, even in increased and highly variable traffic conditions.

Further work will test in detail the performance of QOAS if deployed in local broadband multi-service IP networks against different types of individual traffic flows such as long-lived or short-lived TCP. These tests will study not only the effect that this traffic has on multimedia streams transmitted using QOAS, but also the effect that QOAS streaming has on the other traffic. In this context QOAS's degree of TCP friendliness is of significant importance. Also experiments that involve streaming of more than one type of multimedia clip at the same time are envisaged. Next, QOAS will be extended for multicast transmissions, taking into account some multicast specific characteristics such as multiple feedback and arbitration of heterogeneous client reports in order to make more efficient live multimedia streaming.

ACKNOWLEDGEMENTS

The support of the Research Innovation Fund and Informatics Research Initiative of Enterprise Ireland is gratefully acknowledged.

REFERENCES

- [1] S. Dravida, D. Gupta, S. Nanda, K. Rege, J. Strombosky, M. Tandon, "Broadband Access over Cable for Next-Generation Services: A Distributed Switch Architecture", *IEEE Communications Magazine*, vol. 40, no. 8, Aug. 2002, pp. 116–124
- [2] S. A. Barnett, G. J. Anido, "A Cost Comparison of Distributed and Centralized Approaches to Video-on-Demand", *IEEE JSAC*, vol. 14, no. 6, Aug 1996, pp. 1173-1183
- [3] G.-M. Muntean, L. Murphy, "An Adaptive Mechanism For Pre-recorded Multimedia Streaming Based On Traffic Conditions", Proc. 11th W3C WWW Conf., Honolulu, Hawaii, USA, May 2002
- [4] G.-M. Muntean, L. Murphy, "Adaptive Pre-recorded Multimedia Streaming", Proc. IEEE GLOBECOM, Taipei, Taiwan, Nov. 2002, pp. 1728-1732
- [5] G.-M. Muntean, P. Perry, L. Murphy, "A New Adaptive Multimedia Streaming System for All-IP Multi-Service Networks", *IEEE Transactions on Broadcasting*, vol. 50, no. 1, Mar. 2004
- [6] G.-M. Muntean, P. Perry, L. Murphy, "Performance Comparison of Local Area Video Streaming Systems", *IEEE Communication Letters*, vol. 8, no. 5, May 2004, pp. 326-328
- [7] G.-M. Muntean, P. Perry, L. Murphy, "A Quality-Orientated Adaptation Scheme for Video-on-Demand", *IEE Electronic Letters*, Vol. 39, No. 23, Nov. 2003, pp. 1689-1690
- [8] G.-M. Muntean, P. Perry, L. Murphy, "Performance Assessment of the Quality-Oriented Adaptation Scheme", *IEEE/IFIP Int. Conf. on Management of Multimedia Networks and Services*, San Diego, California, USA, Oct. 2004
- [9] W. Effelsberg, R. Steinmetz, "Video Compression Techniques", Heidelberg, Germany, *dpunkt-Verlag*, 1998
- [10] B. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview", *RFC* 1633, June 1994, http://www.ietf.org/rfc/rfc1633.txt
- [11] S. Blake et. al, "An Architecture for Differentiated Services", *RFC* 2475, December 1998, http://www.ietf.org/rfc/rfc2475.txt
- [12] E. C. Rosen, A. Viswanathan, R. Callon "MPLS Architecture" RFC 3031, January 2001, http://www.ietf.org/rfc/rfc3031.txt
- [13] A. T. Campbell, G. Coulson, D. Hutchison, "A Quality of Service Architecture", ACM Computer Comms Review, Vol. 24, No. 2, Apr 1994
- [14] D. Hutchison, A. Mauthe, N. Yeadon, "Quality-of-service Architecture: Monitoring and Control of Multimedia Communications", *IEEE Electronics and Comms. Eng. Journal*, Vol. 9, No. 3, 1997, pp. 100-106
- [15] ITU-T Recommendation X.641, "Information Technology Quality of service: Framework", Dec. 1997
- [16] ISO/IEC 13236, "Information Technology Quality of service: Framework", International Standards Organisation, December 1997
- [17] A. Banerjea, D. Ferrrari, B. A. Mah, M. Moran, D. C. Verma, H. Zhang, "The Tenet Real-time Protocol Suite: Design, Implementation, and Experiences", *IEEE/ACM Trans. on Networking*, Vol. 4, No. 1, Feb. 1996, pp. 1-10
- [18] L. C. Wolf, R. G. Herrtwich, "The System Architecture of the Heidelberg Transport System", ACM Operating System Review, Vol. 28, No. 2, April 1994
- [19] X. Wang, H. Schulzrinne, "Comparison of Adaptive Internet Multimedia Applications", *IEICE Trans. on Communications*, vol. E82-B/6, June 1999, pp. 806 – 818
- [20] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, J. M. Peha, "Streaming Video over the Internet: Approaches and Directions", *IEEE Trans. On Circuits* and Systems for Video Tech., vol. 11, no. 3, Mar. 2001, pp. 282–300
- [21] D. Wu, Y. T. Hou, Y.-Q. Zhang, "Transporting Real-time Video over the Internet: Challenges and Approaches", *Proc. IEEE*, vol. 88, no. 12, Dec. 2000
- [22] H. Kanakia, P. Mishra, A. Reibman "An Adaptive Congestion Control Scheme for Real-time Packet Video Transport", *Proc. ACM SIGCOMM*, San Francisco, USA, Sep. 1993, pp. 20-31
- [23] J.-C. Bolot, T. Turletti, "A Rate Control Mechanism for Packet Video in the Internet", *Proc. IEEE INFOCOM*, Canada, 1994, pp. 1216-1223
- [24] J. Padhye, J. Kurose, D. Towsley, R. Koodli, "A Model Based TCP Friendly Rate Control Protocol", *Proc. ACM NOSSDAV*, New Jersey, USA, 1999
- [25] J. Padhye, V. Firoiu, D. Towsley, J. Kurose, "Modeling TCP Throughput: A Simple Model and its Empirical Validation", *Proc. ACM SIGCOMM*, Vancouver, Canada, Oct. 1998
- [26] E. H.-k. Wu, C.-S. W. Tseng, C.-Y. K. Chang, M.-f. Tsai, C. R. Lin, "TMRC: A Load-Adaptive TCP-Friendly Rate Control Protocol For

Real-time Multimedia Environment", Proc. IEEE Int. Conference on Comms, Paris, France, June 2004

- [27] S.-Y. Chen, E. H.-k. Wu, Y.-S. Chen, "A New Approach Using Time-Based Model for TCP-Friendly Rate Estimation", *Proc. IEEE Int. Conference on Comms*, Anchorage, Alaska, USA, June 2003
- [28] D. Sisalem, H. Schulzrinne, "The Loss-Delay Adjustment Algorithm: A TCP-friendly Adaptation Scheme", Proc. ACM NOSSDAV, UK, July 1998
- [29] D. Sisalem, A. Wolisz, "LDA+ TCP-Friendly Adaptation: A Measurement and Comparison Study", Proc. ACM NOSSDAV, USA, 2000
- [30] R. Rejaie, M. Handley, D. Estrin, "RAP: An End-to-end Rate-based Congestion Control Mechanism for Realtime Streams in the Internet", *Proc. IEEE INFOCOM*, 1999, pp. 1337-1345
- [31] R. Rejaie, M. Handley, D. Estrin, "Layered Quality Adaptation for Internet Video Streaming", *IEEE JSAC, Special Issue on Internet QoS*, vol. 18, no. 12, Dec. 2000, pp. 2530-2543
- [32] S. Jacobs, A. Eleftheriadis, "Streaming Video Using Dynamic Rate Shaping and TCP Congestion Control", *Journal of Visual Comm. and Image Representation*, vol. 9, no. 3, Sep. 1998, pp. 221-222
- [33] S. McCanne, V. Jacobson, M. Vetterli, "Receiver-Driven Layered Multicast", Proc. ACM SIGCOMM, USA, Aug. 1996, pp. 117-130
- [34] L. Vicisano, J. Crowcroft, L. Rizzo, "TCP-like Congestion Control for Layered Multicast Data Transfer", *Proc. IEEE INFOCOM*, vol. 3, Mar. 1998, pp. 996 – 1003
- [35] I. Rhee, V. Ozdemir, Y. Yi, "TEAR: TCP Emulation at Receivers Flow Control for Multimedia Streaming", Technical Report, CS Department, NCSU, April 2000
- [36] N. Yeadon, F. García, D. Hutchison, D. Shepherd, "Filters: QoS Support Mechanisms for Multipeer Communications", *IEEE JSAC*, vol. 14, no. 7, Sep. 1996, pp. 1245-1262
- [37] L. Wang, A. Luthra, B. Eifrig, "Rate Control for MPEG Transcoders", *IEEE Trans. on Circuits & Systems Video Technology*, vol. 11, no. 2, Feb. 2001
- [38] RealNetworks, SureStream, [Online]. Available: http://www.realnetworks.com
- [39] Microsoft, Windows Media, MBR, [Online]. Available: http://www.microsoft.com
- [40] S. Winkler, A. Sharma, D. McNally, "Perceptual Video Quality and Blockiness Metrics for Multimedia Streaming Applications", Proc. of the International Symposium on Wireless Personal Multimedia Communications, Aalborg, Denmark, September 2001, pp. 553-556
- [41] D. Miras, "Network QoS Needs of Advanced Internet Applications A Survey", Internet2 QoS Working Group, 2002
- [42] C. J. van den Branden Lambrecht, "Perceptual Models and Architectures for Video Coding Applications, *Ph.D. Thesis*, L'Ecole Polytechnique Federale de Lausanne, Switzerland, 1996
- [43] P. Frossard, "Robust and Multiresolution Video Delivery: From H.26x to Matching Pursuit Based Technologies", *Ph.D. Thesis*, L'Ecole Polytechnique Federale de Lausanne, Switzerland, 2001
- [44] Video Quality Experts Group (VQEG), *Final Report*, April 2000, http://www.its.bldrdoc.gov/vqeg/pdf/final_report_april00.pdf
- [45] M. Knee, "The Picture Appraisal Rating (PAR) a Single-Ended Picture Quality Measure for MPEG-2", White Paper, Snell & Wilcox, Jan. 2000, http://www.snellwilcox.com/products/mosalina/content/downloads/parpaper.pdf
- [46] O. Verscheure, P. Frossard, M. Hamdi, "User-Oriented QoS Analysis in MPEG-2 Video Delivery", *Journal of Real-Time Imaging*, vol. 5, no. 5, October 1999, pp. 305-314
- [47] "A Guide to Maintaining Video Quality of Service for Digital Television Programs", White Paper, Tektronix, 2000, http://www.broadcastpapers.com/tvtran/25W_14000_0.pdf
- [48] "Pixelmetrix and KDD Media to jointly market VP Series Picture Quality Analyzer", *Pixelmetrix Press Release*, http://www.pixelmetrix.com/rel/press%20release/1kdd.pdf
- [49] A. B. Watson, J. Hu, J. F. III. McGowan, "Digital Video Quality Metric Based on Human Vision", *Journal of Electronic Imaging*, Vol. 10, No. 1, 2001, pp. 20–29
- [50] A. A. Webster et al., "An Objective Video Quality Assessment System Based on Human Perception" SPIE Human Vision, Visual Processing, and Digital Display IV, Vol. 1913, Feb. 1993, San Jose, USA, pp. 15-26
- [51] A. P. Hekstra et. al, "PVQM A Perceptual Video Quality Measure", Journal of Signal Processing: Image Communication, Vol. 17, No. 10, 2002, pp. 781-798

- [52] S. Winkler "A Perceptual Distortion Metric for Digital Color Video", Proc. of Human Vision and Electronic Imaging SPIE, Vol. 3644, San Jose, USA, January 1999
- [53] C. J. van den Branden Lambrecht, O. Verscheure, "Perceptual Quality Measure Using a Spatio-Temporal Model of the Human Visual System", *Proc. of the SPIE*, Vol. 2668, San Jose, USA, Feb. 1996, pp. 450-461
- [54] A. B. Watson, "Method and Apparatus for Evaluating the Visual Quality of Processed Digital Video Sequences", U.S. Patent No. 6,493,023, Dec. 2002
- [55] S. Wolf, M. H. Pinson, "In-Service Video Quality Measurement System Utilizing an Arbitrary Bandwidth Ancillary Data Channel", U.S. Patent No. 6,496,221, Dec. 2002
- [56] ITU-T Recommendation P.861, "Objective Quality Measurement of Telephone-band (300 - 3400 Hz) Speech Codecs", February 1996
- [57] ITU-R BT.500, "Methodology for the Subjective Assessment of the Quality of Television Pictures", 1974-2002
- [58] ITU-T Recommendation P.910, "Subjective Video Quality Assessment Methods for Multimedia Applications", Sep. 1999
- [59] ITU-T Recommendation P.911, "Subjective Audiovisual Quality Assessment Methods for Multimedia Applications", 1998
- [60] ITU-T Recommendation P.800, "Methods for Subjective Determination of Transmission Quality", Aug. 1996
- [61] ITU-T Recommendation P.920, "Interactive test Methods for Audiovisual Communications", Aug. 1996
- [62] ITU-T Recommendation P.930, "Principles of a Reference Impairment System for Video", Aug. 1996
- [63] G. Ghinea, J. P. Thomas, "QoS Impact on User Perception and Understanding of Multimedia Video Clips", Proc. ACM Multimedia, Bristol, UK, 1998
- [64] NS-2, [Online]. Available: http://www.isi.edu/nsnam/ns/
- [65] NIST Net, [Online]. Available: http://snad.ncsl.nist.gov/itg/nistnet
- [66] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", *RFC 1889*, January 1996, http://www.ietf.org/rfc/rfc1889.txt

AUTHORS' BIOGRAPHIES



Dr. Gabriel-Miro Muntean is a Lecturer with the School of Electronic Engineering, Dublin City University, Ireland, where he obtained his Ph.D. degree in 2003 for research on quality-oriented adaptation schemes for multimedia streaming. He was awarded the B.Eng. and Masters degrees in Software Engineering from the Computer Science Department, "Politehnica" University of Timisoara, Romania. Dr. Muntean's research interests

include QoS and performance-related issues of adaptive solutions for multimedia delivery. He is an IEEE Member.



Dr. Philip Perry (M '92) obtained his Ph.D. in Microwave Engineering from the Department of Electronics and Electrical Engineering at University College Dublin, Dublin, Ireland in 1998. He studied for his Master's degree at the University of Bradford, Yorkshire, England (1989) while his primary degree is from the University of Strathclyde, Glasgow, Scotland. His current research interests are focused on the

applications and enabling technologies for mobile systems.



Dr. Liam Murphy obtained a B.E. in Electrical Engineering from University College Dublin in 1985, and Masters and Ph.D. degrees in Electrical Engineering and Computer Sciences from the University of California, Berkeley, in 1988 and 1992 respectively. He is currently a Senior Lecturer in Computer Science at University College Dublin and the Director of UCD's Performance Engineering Laboratory. His current

research interests include performance issues in multimedia transmission, Voice over IP, and component oriented software systems. Dr. Murphy is an IEEE Member.