Network Performance Criteria for Telecommunication Traffic Types driven by Quality of Experience

Ines Ramadža, Vesna Pekić, and Julije Ožegović

Abstract—A common reason for changing the chosen service provider is user perception of service. Quality of Experience (QoE) describes the end user's perception of service while using it. A frequent cause of QoE degradation is inadequate traffic routing where, other than throughput, selected routes do not satisfy minimum network requirements for the given service or services. In order to enable QoE-driven routing, per-traffic-type defined routing criteria are required. For the purpose of identifying services of interest, we analyzed traffic within a telecom operator network. Next, we defined testbed measurements that explored the impact of packet loss and delay on user QoE for video, voice, and management traffic. For video services, we performed separate measurements for multicast delivery, unicast HTTP Live Streaming (HLS), and unicast Real Time Streaming Protocol (RTSP) traffic. Applying a threshold to QoE values, from the measured dependencies we extracted minimum network performance criteria for the investigated different types of traffic. Finally, we define relevant service classes, for relevant services, we propose the retention or correction of QoE/QOS criteria defined years ago to correspond to traffic scenarios in modern telecom operator networks, and we propose their traffic class priorities.

Index Terms—IP/MPLS routing, minimum network requirements, network performance criteria, network traffic measurements, Quality of Experience (QoE), Quality of Service (QoS), telecommunication services, traffic analysis

I. INTRODUCTION

The amount of traffic in telecommunication networks is continuously growing, and users are becoming more demanding. Service providers are attempting to maximize the utilization of available network resources, while at the same time trying to ensure the necessary level of service quality. Internet Protocol/Multiprotocol Label Switching (IP/MPLS) technology, and employing MPLS tunnels for network virtualization provides tools for increasing the level of network quality and availability in today's telecom networks. The same network infrastructure transfers a large number of different video, data, and voice services, requiring different treatment when routed through an IP/MPLS network.

Methods of traffic analysis enable quality dimensioning, the planning of necessary resources, designing independent redundancy paths, and ensuring network availability in case of partial capacity disruption. Problems occur when there is enough capacity, but the quality of certain paths at a certain moment is not satisfactory for one or more services. To adequately perform traffic routing, it is necessary to be aware of the minimum network performance criteria for each traffic type. This study is part of our research in the field of Quality of Experience (QoE) provisioning in MPLS networks [1],[2].

Network performance objectives for different traffic types have been provided in ITU-T Recommendations and adopted by other standardization bodies. The most recent version of the relevant recommendations was updated in 2011. It notes the emergence of new applications, and that their performance objectives require further study. Other sources are based on said ITU Recommendations, or alternatively, contribute few improvements. One way to obtain the minimum network performance criteria is to first determine the dependency of user QoE on network parameters. Numerous efforts to determine this relationship have produced a number of objective models for different traffic types, depending on different parameters. Although some of the models proposed in the literature can be adjusted and applied to obtain the required dependency for the individual types of traffic we investigated, we opted for a different approach. To achieve more unified results, we derived our own measurements. We included monitoring/management traffic, as well as HTTP adaptive bitrate streaming traffic, to the best of our knowledge not covered elsewhere.

The methodology used to obtain the required network performance criteria is shown in Fig. 1. Source signals were subjected to emulated parametric impairments (packet loss, delay, and delay variation) while passing through the network. For video services, the quality of source signal was controlled to ensure accurate quality rating results. The received signal was evaluated using tools based on objective QoE media layer...
(perceptual) models, and QoE was estimated. Perceptual models served as our tool of choice, because they are specifically designed to emulate subjective quality ratings. For management traffic, subjective QoE estimation was performed. Mean Opinion Score (MOS) scale [3, p. 1] was used for QoE results, as MOS is considered relatively well-suited for use by service providers and telecom operators for quality monitoring and alerts [4]. From the estimated QoE and corresponding parameter values, we obtained the dependency of QoE level on the chosen QoS parameter. After applying a threshold QoE MOS value, we extracted the criteria for a minimum level of network performance that would ensure satisfactory user QoE.

The main contribution of this work is the redefinition of relevant service classes, the re-evaluation and extension of the available thresholds in order to ensure that they are applicable as QoE-driven network performance criteria for contemporary telecom network services, and the assignment of traffic type (traffic class) priorities based on the observed thresholds. We re-evaluated the criteria for control traffic, voice over IP (VoIP) and video, and extended the latter with separate criteria for multicast, unicast Real Time Service Protocol (RTSP), and unicast HTTP Live Streaming (HLS) traffic. The results were compared in detail with those available in ITU-T Recommendations and other sources. We also present concise results of a user traffic analysis we conducted in a telecom operator core network, in order to identify services of interest.

The remainder of the paper is organized as follows: Section II provides an overview of related works. Section III first provides brief results of the observed telecom operator user traffic analysis. Following on, it describes the testing environment for video services, and the measured impact of network performance for multicast and unicast video traffic. A description of the testing environment and measured impact of network performance for voice services follows. Finally, the section presents measurements of network performance impact on monitoring (management) traffic, and comments on internet and data services. Summarized measurement results and extracted network performance criteria are presented in Section IV. At the end of the section, obtained results are compared with results from sources in the bibliography. Conclusions are provided in Section V.

II. RELATED WORKS

QoE has been a topic of focus in both industry and academia in recent years. Although there exist standardized QoE models ([5], [6]), the exact definition of QoE remains under discussion [7], [8]. Knowing how QoS parameters can affect user QoE is an important issue for improving telecommunication services.

Quantified network performance objectives for achieving appropriate QoS/QoE for different types of traffic are provided in ITU-T Recommendations. ITU-T recommended performance targets for several audio, video, and data applications at the start of the 2000s [9]. In [10], guidance for one-way delay is provided, with special attention paid to achieving satisfactory delay with VoIP. A more recent document [11] offers provisional values for the minimum level of transport layer performance required to provide satisfactory IPTV QoE. Nevertheless, it also states that the exact criteria may vary according to the requirements defined by each IPTV service context, thereby recognizing video delivery technique as an influencing factor for video services. IP network QoS class definitions and network performance objectives are discussed in [12].

Based on ITU-T recommendations, [13] describes QoS requirements for multimedia services, according to four service classes defined in satellite networks. Additionally, experimental results of network QoS requirements from a Next Generation Network testbed for representatives of audio streaming, games, and VoIP services are presented. In [14], a proposal is made for a cheat-proof framework for measuring QoS that users find intolerable, thus obtaining minimum QoS needs for real-time networked multimedia services. Network bandwidth and loss rate intolerance thresholds for several VoIP, video conferencing, and network gaming applications are also estimated.

Several surveys can help to gain insight into the field of QoE assessment. Correlation models mapping QoS to QoE for multimedia services are reviewed in [15]. A survey on QoE of HTTP Adaptive Streaming [16] considers perceptual and application layer QoE influence factors, as well as interaction with the TCP control loop. A study of tools and methodologies for assessing the QoE of online Video on Demand (VoD) services is presented in [17]. A comprehensive survey of video
QoE assessment methods is given in [18], referencing subjective, objective, and data-driven assessment. A study of VoIP QoE evaluation approaches is detailed in [19]. In [20], details are provided of parametric QoE models for a range of popular voice, video, and data services. With additional to presenting models from the literature, it describes standardized ITU-T parametric network planning models for VoIP and online videophone applications, and gives their numerical results; in particular, the dependency of VoIP QoE (MOS) on delay and on packet loss, and the dependency of video QoE on audio/video delay, on packet loss, and on video frame rate requirements.

A wide range of parameters influence QoE, and modeling their influence is highly service dependent. We focused on works that considered the correlation between QoE and network level parameters such as packet loss, delay, and delay variation in wired networks.

The correlation of QoE (measured using Opinion Score) and parameters relevant for internet/data service was studied in [21]. The results demonstrate an exponential relationship between QoE and download time, a linear relationship between QoE and nominal loss, and a logarithmic relationship between QoE and throughput.

Several papers have explored machine learning approaches for determining QoE. In [22], the efficacy of several machine learning methods is reviewed for assessing the QoS/QoE relationship for video streaming service. Predicting VoIP QoE based on latency, jitter, and packet loss, using a framework with modular machine-learning algorithms, is studied in [23].

A generic, exponential relationship between a single QoS parameter and QoE is proposed in [24], and validated using several case studies in [24] and [25]. Among others, an exponential mapping function was found between packet loss ratio and QoE for the iLBC and G.711 VoIP codecs. Extending [24], [26] provides an exponential QoS/QoE correlation model, where QoS is modeled as a weighted sum of several parameters. VoD (Video on Demand) QoE was taken as a use case for validation. First, QoV (Quality of Video), depending on packet loss rate and delay, was measured using a subjective method and the Opinion Score used in [21]. Following on, a composite exponential correlation function for QoV was obtained.

In [27], a power relationship is proposed between video streaming QoE and packet loss, and the model is validated using the statistical analysis of data obtained with the objective peak signal to noise ratio algorithm (PSNR) [28].

A model for IPTV video QoE presented in [29], derives QoE from a normalized QoS value calculated, again, as a weighted sum of several parameters (e.g., packet loss, burst level, packet jitter, packet delay, and bandwidth), exact parameters and weights depending on the service.

In [30], QoE is modeled as a logarithmic utility function of the considered application, and weighted user-level Key Quality Indicators (KQIs). KQIs are themselves dependent on the application and network level parameters. Logarithmic relationships can be found between certain network parameters and QoE for specific services, e.g., downlink bandwidth or connection setup time for web browsing.

III. IMPACT OF NETWORK PERFORMANCE ON TELECOMMUNICATION SERVICES QUALITY

A. Analysis of User Traffic in the Core Network

The goal of network traffic analysis in a Next Generation Network (NGN) is to recognize the daily, weekly, monthly, and yearly patterns in user behavior for each type of traffic, in order to detect growth trends, and to determine the share of each traffic type in the total amount of traffic. This information enables tailor-made network planning and routing.

We analyzed user traffic patterns and characteristics from data collected over a period of one year, from a group of users connected to the observed telecom operator network, using both narrowband and wideband technologies. SNMP (Simple Network Management Protocol) was used to collect data from network devices every 30 seconds. Points of measurement in

| Traffic | Total | VoIP | POTS | IPTV | Internet |
|---------|-------|------|------|------|----------|
| Max IN (Mb/s) | 619.45 | 6.52 | 2.95 | 182.41 | 427.57 |
| Max OUT (Mb/s) | 57.56 | 7.34 | 3.38 | 0.06 | 46.78 |
| Share IN (%) | 100 | 1.05 | 0.48 | 29.45 | 69.02 |
| Share OUT(%) | 100 | 12.75 | 5.87 | 0.10 | 81.27 |
the network varied in order to take into account relevant groups of users for each traffic class. The number of users varied, depending on the analysis, from several hundred to several tens of thousands.

After analyzing different user traffic categories, the following traffic usage and characteristics were identified:

1) Internet and video traffic were the dominant categories for total traffic within the telecom operator’s autonomous system (see Table I and Fig. 2), with 69.02% and 29.45% participation in the overall traffic of the transit route, respectively.

2) The amount of traffic used for network and user equipment monitoring was negligible with respect to the overall traffic in the network. The daily maximum for monitoring approximately 1300 user devices in a selected geographic area, as well as related core network devices, required 54.31kb/s in, and 28.27kb/s out capacity.

3) Internet and video traffic are continuously growing, while voice traffic is decreasing.

4) Occasity of transit routes is asymmetric, with more than 10 times greater occupancy from the network to the user.

5) All traffic categories had predictable temporal patterns, making it possible to anticipate the required capacities in the core of the network.

6) Business users had different hours of daily peak load for voice and internet services, compared to residential users; however, business users’ traffic patterns were also predictable.

In Table I and Fig. 2, we present user traffic volume and percentage composition by service type, measured on a transit route towards the rest of the MPLS network over the yearly timescale.

For the measurements described in the following subsections, a total of 22 WAN degradation scenarios were used: Test0 as a no loss no delay reference, Tests 1-11 for paths with packet losses, and Tests 12-21 for paths with delay and jitter.

**B. Impact on Video Services Quality**

The analysis of network traffic showed that video traffic comprised a considerable share in the total amount of traffic. Therefore, it is necessary to explore the criteria of minimum network performance required to keep QoE at a satisfactory level. Since the delivery of video services relies on both multicast and unicast traffic, which demand different requirements from the network, the impact of network performance was tested on three types of traffic:

1) multicast traffic used to deliver a program in real time
2) unicast HTTP Live Streaming (HLS) [31] traffic, used mostly in over-the-top (OTT) delivery
3) unicast Real Time Service Protocol (RTSP) [32] traffic, which can be used for delivery of video-on-demand (VoD) time-shifting content

**1) Testing environment (testbed)**

The testbed for testing and analyzing network performance impact on QoE for video flows was put together using the following:

**--Internet Protocol television (IPTV) Headend system**
**--Protocol Independent Multicast (PIM) router [33]**
**--Wowza streaming engine server [34]**
**--Edgeware Video Consolidation Platform [35]**
**--Wide Area Network (WAN) connection emulator [36]**

**--AccepTV MPEG (Moving Pictures Experts Group) Monitor / Video Quality Monitor [37]**

A logical view of the testbed is presented in Fig. 3.

An IPTV Headend system receives the TV signal from a satellite, codes the signal into MPEG-2 or MPEG-4, and multicasts it to the IP network. While testing multicast traffic flows, two programs were used in real time to eliminate the effect that error on the source might have on the analysis of results. Quality of the source signal was monitored using two quality control tools: Bridgetech probes [38] and MPEG Monitor.

The Edgeware Video Consolidation Platform sends unicast audio and video flows using RTSP. RTSP allows control over video data flow, viewing options, and pausing and stopping content. The content from [39] was used for testing.

The Wowza streaming engine was employed for sending unicast audio and video flows using the HLS protocol. The content from [40] was used for testing.

The AccepTV MPEG Monitor software package for video and audio quality control enables perceived quality (QoE) monitoring in real time, with the aid of Human Vision System model based on [41]. It displays video and audio quality for each flow, as well as video and audio data rate. In addition to MPEG Monitor, AccepTV Video Quality Monitor (VQM) was used for the measurement and monitoring of end user audio

---

| Video audio quality | MOS  |
|---------------------|------|
| 80 - 100            | Excellent (5) |
| 60 - 80             | Good (4)       |
| 40 - 60             | Fair (3)       |
| 20 - 40             | Poor (2)       |
| 0 - 20              | Bad (1)        |
and video quality perception (QoE). VQM expresses QoE in MOS scale. The video quality metric is optimized to generate quality results that are highly correlated with human signal quality assessment. The audio and video quality Mean Opinion Score reference scale is shown in Table II.

SoftPerfect [36] Connection Emulator emulates the WAN environment. It was used in the testbed to simulate the degradation of network performance.

2) Multicast traffic

For multicast video transmission two groups of tests were conducted. The first simulated the percentage of lost packets on the transmission path, while the other simulated the increase in delay and delay variation on all, or on a certain number of packets. All testing was conducted on the testbed shown in Fig. 3, using two different real-time program flows, received from two different satellites, using different transcoders (see Table III.) The results shown in the paper are those related to multicast flow for IP address 233.121.36.170. The other multicast flow was used for results control. Quality control of the source signal was performed for the entire duration of testing. Quality of the received signal prior to the impact of the WAN emulator was evaluated with MOS rating 5 (perceived video quality between 80 and 100).

The first group of tests simulated a certain percentage of lost packets in the network. Preliminary testing showed that multicast video flow was very sensitive to packet loss; 0.1% lost packets caused transmission errors resulting in periodical degradation of video signal quality, perceived by the control tool, as well as visible in the decoded video signal (see Fig. 4). The worst achieved quality was 58.09, corresponding to MOS rate 3. The increase of emulated packet loss rate resulted in a further decrease of MOS. At 0.3% lost packets, the lowest value of video quality decreased to 36.27, corresponding to MOS 2. To summarize, increasing the packet loss rate from 0.3% increased the frequency of visible error in the picture, while MOS was lowered from 5 to 2. An overview of tests and results is provided in Table IV.

The second group of tests examined the impact of delay and delay variation on the quality of multicast video transmission. We obtained a range of realistic values for testing delay by tracking baseline values on the links implemented through different technologies. Simulation of realistic values for delay and delay variation revealed no impact on quality. We then conducted a series of tests with delays ranging from realistic values up to values of 5000 ms. A test with 1000 ms of delay for all packets produced no visible impact on video signal quality or MOS. Additional tests introduced a random delay ranging from 1 to 5000 ms on 100% of packets. Then measured the impact on MOS in case of random delay, ranging from 100 to 1000 ms, on 50% of packets. Thus, we simulated flow transmission using different network paths, with one path having considerable delay. Delays introduced randomly on 50% of packets also meant that delay variation was introduced. In order to try more aggressive degradation of delay and delay variation, we conducted additional tests with a broader range of delays, as well as different percentages of influenced packets. However, none of the tests revealed an impact on QoE (Table IX).

Delay and delay variation had no impact on video quality due to buffers implemented on STB (Set-Top Box) devices, precisely for the purpose of eliminating delay and delay variation.

We can therefore conclude that, in the case of multicast video services delivery, lost packets have the most significant impact on service quality. Increasing the percentage of lost packets leads to an increase in errors in the picture, and a decrease in MOS. To achieve as high level of QoE as possible, multicast video traffic should be directed to routes where either there are no transmission errors, or the ratio of lost packets in the total number of packets is as low as possible (lower than 0.1%).

3) Unicast HLS traffic

In the testbed (Fig. 3), the source of unicast traffic is the Wowza server, accessible on the internet via a public IP address. Prior to and during testing, video signal quality was controlled by MPEG Monitor. The quality of the observed content on the source prior to and during testing was evaluated with the highest MOS.
In the first group of tests, simulating lost packets of flow under observation, tests with low packet loss rates indicated no influence on flow quality. Unlike User Datagram Protocol (UDP) multicast flows, HLS protocol employs HTTP adaptive streaming based on the Transmission Control Protocol (TCP). The network packet analyzer revealed the retransmissions of lost packets. These retransmissions, along with initial buffering, allowed for no visible impact on the service.

Gradual increase in the percentage of lost packets eventually led to frame freezing and jerkiness, starting at the 9% level. For this level, measurements showed a drastic drop in MOS (see Table V). MPEG Monitor and Video Quality Monitor reported the lowest perceived video quality of 20.96, corresponding to the lower edge value of MOS 2. The level of 9% lost packets is therefore a critical limit over which retransmissions cannot be performed in time, and QoE becomes drastically impaired.

The second group of tests, simulating delay and delay variation, showed no impact of delay and delay variation on the quality of HLS flow, and no visible degradation of video signal. With HLS, at least one stream segment containing roughly 10 seconds of media content was first downloaded from the server, prior to being reproduced on the client side. Therefore, even large values of delay/delay variation had no impact on QoE.

The results for both groups of tests are shown in Table IX. We can conclude that HLS flows are much more resilient to the degradation of network performance than multicast flows. Increased delay and/or delay variation had no impact on quality, and lost packets caused frame freezing and jerkiness that only started at a high packet loss rate of 9%.

4) Unicast RTSP traffic

Unlike HLS, which essentially uses TCP, RTSP uses UDP and does not provide for the possibility of checking packet delivery status and retransmitting in the case of unsuccessful delivery.

RTSP traffic testing was performed with the help of the Edgeware system used for VoD services, as shown in Fig. 3 (testbed). The quality of the observed content on the source prior to and during testing was evaluated with the highest MOS rating. Quality of RTSP flow prior to the impact of the WAN emulator, shown in Table VI, was evaluated with MOS rating 5. The worst measured value of video quality in the observed period was 77.07, which is close to the upper limit of MOS 4, and there was no visible impact on the service quality.

The first group of tests, simulating lost packets, produced results expected for UDP based flows. As was the case with multicast flows, quality degradation was already noticed at 0.1% of lost packets. Artifacts in the picture were also visible. Video signal quality was lowered to MOS 3. Increasing the percentage of lost packets caused increased error frequency and decreased MOS.

The second group of tests, simulating delay and delay variation, showed no impact of realistic values of delay and delay variation on RTSP flow. Further testing was done with different values of delay and delay variation on all, or on a certain percentage of packets. The tests revealed no impact on quality.

Results for both groups of tests are shown in Table IX. We can thus conclude that packet loss had the greatest influence on the quality of RTSP video services. Similar to multicast traffic, increasing the percentage of lost packets led to an increase of errors in the picture and a decrease in MOS. To achieve as high a level of QoE as possible, RTSP video traffic should be directed to routes where either there are no transmission errors, or the ratio of lost packets in the total number of packets is as low as possible (again, lower than 0.1%).

C. Impact on Voice Services (VoIP) Quality

1) Testing environment

The testbed for testing minimal network performance for voice traffic (shown in Fig. 5) included:

--SIP (Session Initiation Protocol [42]) PBX (Private Branch Exchange) and a SIP trunk connected to a public voice service
--WAN connection emulator [36]
--Network analyzer (OmniPeek [43])
--Softphone client As SIP PBX, we used Fusion PBX [44], an open source FreeSwitch employable as multi-tenant PBX, call center server, and VoIP server. In the testbed, Fusion PBX was used as a SIP proxy server for the registration of the SIP softphone client. Fusion PBX was connected with a SIP trunk to the Class4/5 softswitch in the testbed to enable communication between the softphone client and the public telecommunications network.

![Fig. 5. Testbed for voice services.](image-url)

| Video measurements | Current value | Worst value |
|--------------------|--------------|-------------|
| Perceived video quality (MOS) | 83.61 | 77.07 |
| Blockiness | 0.60 | 1.21 |
| Blur | 1.46 | 1.95 |
| Jerkiness | 0.0 ms | 0.0 ms |
SoftPerfect Connection Emulator is a WAN environment emulation software package. It was used in the testbed for the simulation of network performance degradation, allowing degradation in one or both directions, bandwidth limiting, introducing packet loss and delay in traffic flows, duplicating, and reordering.

OmniPeek Network Analyzer was used for the analysis of voice traffic quality. OmniPeek provides VoIP RTP flow and signaling analysis, MOS and R-Factor [5] results, detailed visual flow of each call, and the possibility of replaying a captured call. The correlation of MOS values, R-Factor values, and the perceived call quality is given in [5], Annex B.

Since capacity no longer poses problems, codec G.711 is employed in many telecom networks. It has been used exclusively in the observed telecom network, and consequently, in the testbed. Codecs with greater compression, such as G.729 have a smaller tolerance for lost packets than G.711; this needs to be taken into account when using the results of this study. The measured starting reference values of call quality using G.711 are given in Table VII (Test 0). The call was rated with R-Factor 93 (very satisfied), as was expected according to the reference table.

2) Measured impact

The first group of tests measured the impact of lost packets on call quality. The results, provided in Table VII, revealed that MOS was lowered to 4.03 and R-Factor to 87/86 (listening/conversational), which was already at 0.5% of lost packets (Test 1). The quality of the observed call was still in the category where most of the users would be satisfied. Test 2, simulating 0.7% of lost packets, yielded a MOS value of 3.91 and R-Factor values of 82/81. A call with 0.7% lost packets is on the border between satisfactory and unsatisfactory for a certain number of users. In Test 3, a call with 1% of lost packets resulted in MOS 2.91 and R-Factor 60/59, making it a call that would leave many users dissatisfied. In Test 4, the percentage of lost packets was increased to 2%, while quality ratings, as expected, decreased. MOS value 2.66 and R-Factor values of 55/54 put the call in the category where almost all users would be dissatisfied. To conclude, packet loss rates below 0.5% had no significant impact on call quality, while packet loss rates between 0.5% and 2% did.

The second group of tests measured the influence of delay on call quality. Test results are shown in Table VII. By introducing delays of 150 ms, Test 5 produced a MOS value of 3.65 and R-Factor values of 75/78. These rates fall into the category where the quality of the call for some users will be satisfactory, but is nonetheless very close to the lower border. Test 6 simulated a 250 ms delay. MOS and R-Factor were rated 3.52 and 72/73, respectively, indicating that many users would be dissatisfied. Test 7 simulated a 350 ms delay on 100% of packets. This call received very low rates: MOS value was 2.02 and R-Factor values 68/41, describing a call of very low quality (not recommended).

Table VII

| Parameters and measured quality | Test 0 | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 | Test 7 |
|---------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Codec G.711A | baseline | increased packet loss | increased delay |
|----------------|-------|-------|-------|-------|-------|-------|-------|
| Transmission rate (b/s) | 64000 | 64000 | 64000 | 64000 | 64000 | 64000 | 64000 | 64000 |
| Duration (s) | 14.059530 | 19.919362 | 26.159918 | 17.207146 | 11.082090 | 24.440238 | 25.140271 | 13.349552 |
| MOS-Low | 4.17 | 4.03 | 3.91 | 2.91 | 2.66 | 3.65 | 3.52 | 2.02 |
| One-way delay (s) | 0.048 | 0.053 | 0.052 | 0.053 | 0.053 | 0.153 | 0.253 | 0.359 |
| Lost packets (%) | 0 | 0.5 | 0.7 | 1 | 2 | 0 | 0 | 0 |
| Delay variation (s) | 0.000679 | 0 | 0 | 0.026468 | 0 | 0 | 0 | 0 |
| MOS-LQ | 4.19 | 4.06 | 3.91 | 2.96 | 2.71 | 3.69 | 3.56 | 3.34 |
| MOS-CQ | 4.17 | 4.03 | 3.87 | 2.91 | 2.66 | 3.65 | 3.52 | 2.02 |
| MOS-PQ | 4.44 | 3.93 | 3.80 | 3.29 | 3.14 | 3.70 | 3.63 | 3.51 |
| MOS-Norm | 4.19 | 4.19 | 4.19 | 4.19 | 4.19 | 4.19 | 4.19 | 4.19 |
| R Factor | 93 | 87 | 82 | 60 | 55 | 76 | 73 | 68 |
| Listening | | | | | | | | |
| Conversational | | | | | | | | |
| R Factor G.107 | 92 | 86 | 81 | 59 | 54 | 75 | 72 | 41 |
| R Factor nominal | 93 | 93 | 93 | 93 | 93 | 93 | 93 | 93 |
| Descriptive rating | Very satisfied | Satisfied | Satisfied, Some users satisfied | Many users dissatisfied | Nearly all users dissatisfied | Some users satisfied | Many users dissatisfied | Not recommended |

Table VIII

| Test nr. | Lost packets (%) | Delay (ms) | Result |
|---------|------------------|-----------|--------|
| 1 | 0.5 | -- | Does not impede work |
| 2 | 0.6 | -- | Does not impede work |
| 3 | 0.7 | -- | Perceived during work |
| 4 | 0.8 | -- | Perceived during work |
| 5 | 1 | -- | Impedes work |
| 6 | 1.5 | -- | Impedes work |
| 7 | 2 | -- | Impedes work |
| 8 | -- | 10 | Does not impede work |
| 9 | -- | 15 | Perceived during work |
| 10 | -- | 20 | Perceived during work |
| 11 | -- | 50 | Impedes work |
does not allow for control over the exact value of delay variation. Due to restrictions of the available tool, the value of 1 ms for the criterion of maximum delay variation in Table X was adopted from [9].

D. Impact of Network Performance on Monitoring, Internet and Data Traffic

Analyzing the ratio of different types of traffic in the total transit route traffic revealed that monitoring (management) traffic occupied negligible capacity compared to total network requirements. Telecommunication equipment is most frequently monitored using Telnet [45], SSH [46], SNMP [47], and TR-69 CPE WAN [48] management protocols. These protocols use two-way communication and losing packets hinders their interactive functioning. The same can be said for the impact of communication delays. It is of interest to the operator and, indirectly, to the users, that minimum network requirements are met that will enable smooth monitoring and management. To obtain these requirements, testing was done on Telnet traffic using the subjective user feedback method. The results, provided in Table VIII, show that packet loss greater than 0.6% influenced server-client interaction, as well as a delay of 15 ms or greater.

Since monitoring traffic requires negligible capacities in the core of the telecommunication network, it should be directed to routes with the best virtual path measurement results in the MPLS TE network.

Internet and data services transfer different types of traffic with different needs and characteristics. Due to a broad range of services, the impact of delay, delay variation, and packet loss has not been investigated. Internet service traffic is, from the perspective of traffic volume, the most important traffic category in the service provider network, and frequently classified as the 'best effort' class. Data services traffic, mostly reserved for business users, is described using a category based on the contracted service.

IV. RESULTS AND DISCUSSION

A. Summarized measurements results

The combined results of simulating network performance degradation and measuring its impact on user experience for voice, video, and monitoring traffic are summarized in Table IX. The results can now be used for extracting the starting criteria in the decision of routing the specific protocol traffic, as well as in traffic categorization. Another possible usage is increasing the visibility of network operation or QoE level estimation. Knowing the impact of network performance on QoE, the central system can show the level of QoE in any part of a MPLS network, based on measurements and determined criteria.

B. Criteria for minimum network performance and resulting traffic categories

We selected the threshold value for satisfactory QoE as being between MOS 4 and 3. Based on the results shown in Table IX, the starting criteria for minimum network performance per traffic type were derived. These are listed in

| Test nr. | Lost packets (%) | Delay (ms) | Delay variation (ms) | MOS mult | MOS HLS | MOS RTSP | MOS mgmt | MOS VolP |
|----------|------------------|------------|----------------------|----------|---------|---------|----------|----------|
| 0        | 0                | --         | --                   | 5        | 5       | 5       | 5        | 4        |
| 1        | 0.1              | --         | --                   | 3        | 5       | 3       | 5        | 4        |
| 2        | 0.2              | --         | --                   | 3        | 5       | 2       | 5        | 4        |
| 3        | 0.3              | --         | --                   | 2        | 5       | 2       | 5        | 4        |
| 4        | 0.5              | --         | --                   | 1        | 5       | 1       | 5        | 4        |
| 5        | 0.7              | --         | --                   | 1        | 5       | 1       | 5        | 3        |
| 6        | 1                | --         | --                   | 1        | 5       | 1       | 5        | 2        |
| 7        | 2                | --         | --                   | 1        | 5       | 1       | 5        | 2        |
| 8        | 5                | --         | --                   | 1        | 5       | 1       | 1        | 1        |
| 9        | 8                | --         | --                   | 1        | 5       | 1       | 1        | 1        |
| 10       | 9                | --         | --                   | 1        | 2       | 1       | 1        | 1        |
| 11       | 10               | --         | --                   | 1        | 2       | 1       | 1        | 1        |
| 12       | --               | 10         | --                   | 5        | 5       | 5       | 5        | 4        |
| 13       | --               | 15         | --                   | 5        | 5       | 5       | 3        | 4        |
| 14       | --               | 150        | --                   | 5        | 5       | 5       | 1        | 3        |
| 15       | --               | 200        | --                   | 5        | 5       | 5       | 1        | 3        |
| 16       | --               | 300        | --                   | 5        | 5       | 5       | 1        | 2        |
| 17       | --               | 1000       | --                   | 5        | 5       | 5       | 1        | 1        |
| 18       | --               | 5000       | --                   | 5        | 5       | 5       | 1        | 1        |
| 19       | --               | 100-500    | random               | 5        | 5       | 5       | 1        | 1        |
| 20       | --               | 1-5000     | random               | 5        | 5       | 5       | 1        | 1        |
| 21       | --               | 50% pkts. 100-1000 | random               | 5        | 5       | 5       | 1        | 1        |

TABLE X

| Traffic type                  | Category | Packet loss rate | Delay | Delay variation |
|------------------------------|----------|------------------|-------|-----------------|
| Management / monitoring (Telnet) | 7        | < 0.7%          | < 15 ms | --              |
| Voice (VoIP G.711A)          | 6        | < 1%            | < 150 ms | < 1 ms           |
| Video-multicast              | 5        | < 0.1%          | --    | --              |
| Video-unicast RTSP           | 5        | < 0.1%          | --    | --              |
| Video-unicast HLS            | 4        | < 9%            | --    | --              |
| Internet traffic             | 0        | --              | --    | --              |
Table X, along with the newly assigned traffic category. Traffic category can be used in traffic routing to determine the priority of a traffic type when choosing network paths, and for using up capacity.

Taking into account very strict obtained network performance criteria, and the fact that it occupies negligible capacity, management and monitoring traffic was assigned the highest category. From the measurements analysis, we can conclude that video and voice traffic are very sensitive to deteriorated network performance. Packet loss rates greater than 1%, as well as delays exceeding 150 ms, caused decreased QoE for voice traffic. Since video traffic is not sensitive to increased delay, it is assigned a lower category than voice traffic. It is recommended that video traffic transmitted using different protocols be individually categorized, since protocols react differently to deteriorated network performance. This type of traffic makes up a considerable part of the total telecommunications traffic and can contribute significantly to customer satisfaction. For multicast and RTSP video traffic, transmission without errors must be insured. Therefore, multicast and RTSP video traffic should be in a higher category than the more robust HLS traffic.

Let us observe an example, where management, IPTV, HLS, VoIP and internet/data traffic is to be sent through two MPLS tunnels: tunnel A, characterized by 10 ms delay and 0.005 packet loss rate, and tunnel B, characterized by 300 ms delay and 0.00002 packet loss rate. Taking into account the above criteria and categories, tunnel A will be used for routing management and VoIP traffic. The remaining capacity will be used first for HLS traffic, and then for internet/data traffic. Tunnel B we will be used primarily for routing IPTV traffic, then for any remaining HLS traffic, and, upon availability, for the internet/data traffic.

| Traffic type | Source | Packet loss rate | Delay | Delay variation |
|--------------|--------|------------------|-------|-----------------|
| Management / monitoring (Telnet) | this study | <0.007 | <15 ms | -- |
| Command / control traffic (Telnet) | G.1010 | 0 | <250 ms (200 ms) | -- |
| Voice (VoIP G.711A) | this study | <0.01 | <150 ms | <1 ms |
| Voice (conversational) | G.1010 | 0.03 | <150 ms | <1 ms |
| Voice (VoIP) | Y.1541 | <0.001 | <100 ms (400 ms) | <50 ms |
| Voice (VoIP app.) | Ref. [13] | 0.1 | <150 ms | <50 ms |
| Voice (VoIP app.) | Ref. [14] | 0.085 | -- | 0.125 |
| Video (multicast IPTV) | this study | <0.001 | -- | -- |
| Video - unicast RTSP | this study | <0.001 | -- | -- |
| Video - unicast HLS | this study | 0.9 | -- | -- |
| Video (one-way) | G.1010 | <0.01 | <10 s | -- |
| Video streaming | Y.1541 | <0.001 | <1 s | -- |
| Video (IPTV) | Y.1541 | <10⁻⁴ | <100 ms (400 ms) | <50 ms |
| Video (broadcast IPTV and VoD) | G.1080 | <10⁻⁸ | -- | <50 ms |

C. Comparison with Results from Bibliography Sources

A number of transport layer performance criteria for ensuring satisfactory QoS/QoE are provided in literature listed in Section II. These include ITU-T Recommendations G.1010, G.114, G.1080, and Y.1541, and two additional sources. We present the relevant criteria in Table XI for easier comparison with the criteria we obtained in Table X.

In G.1010, command/control applications like Telnet are characterized as error intolerant and interactive, requiring delays <<1s. Suggested delay performance target is <250ms for general command/control applications, and <200 ms for asymmetric Telnet, while loss target is 0%. These values differ from our obtained criteria for management traffic; however, they are not incompatible. Targeted 0% packet loss rate is, in fact, the allowed information loss rate. Our packet loss rate criterion for Telnet was <0.7%; however, the non-zero packet loss rate is compensated for by packet retransmissions, up to a point where retransmission-induced delay began influencing performance. Our delay criterion is an order of magnitude smaller. We are inclined to attribute the difference to expectations, which have changed in the past 17 years. The delay is a one-way, end-to-end delay, which included delays in the terminal, network, and any server.

According to G.114, it is desirable to keep end-to-end delays (observed by user applications) below 150 ms, and in general, to have network planning not exceed 400 ms, unless it is unavoidable due to long transmission distances. It is recognized that even 150 ms may be too loose a requirement for selected applications (e.g., some interactive traffic, such as management/monitoring traffic).

Network performance objectives for VoIP and video taken from Y.1541 are objectives for the assigned network QoS class, meaning that the values may be more stringent for a particular application. The delays are pure network one-way delays; terminal delays will add from 50 to 80 ms.

VoIP is assigned to class 0 and class 1 in Y.1541. The required loss objective of <10⁻³ for these classes is an order of magnitude smaller than ours. On the other hand, references [13] and [14] offer loss criteria that are an order of magnitude greater. The exact VoIP application used is not mentioned in [13], but results in Table XI from [14] pertain to Google Talk, Skype, AOL Instant Messenger, and MSN Messenger. Packet loss rate performance target in G.1010 is somewhat looser than ours (<3% compared to <1%). It is noted in G.1010 that the loss performance targets for audio and video applications depend on specific codecs, but assume use of packet loss concealment. The criterion for VoIP delay of <150 ms in [13] matches ours exactly, and the one for two-way conversational voice audio from G.1010. Class 0 from Y.1541 requires a <100 ms delay, which also complies with our results when terminal delays are added. The criterion for delay variation in [13] matches that of Y.1541, but is considerably looser than the one in G.1010, which we adopted.

The target for packet loss rate in G.1010 is estimated as being <1%, which is 10 times greater than our obtained values for multicast and RTSP unicast traffic, and smaller than the 9% we obtained for unicast with HLS. However, this value
was not examined in detail at the time and video applications have since evolved significantly. Comments and provisional performance parameters for IPTV in G.1080 suggest that packet loss rate should be as small as possible, with the threshold value of the order of magnitude $10^{-6}$. The exact value depends on transport stream bit rate, codec, maximum duration of an error event, and required loss distance. Our lowest tested impairment level for the matching IPTV traffic packet loss rate (0.1%) already resulted in degraded IPTV, which supports these suggestions. Video streaming is assigned to Class 4 in Y.1541 and the resulting objective for packet loss rate of $<10^{-3}$ matches ours. The objectives for video IPTV were taken from provisional classes 6 and 7, since it was recognized that video services are evolving and require further investigation. The packet loss criterion of $<10^{-5}$ approaches that of G.1080.

In G.1010, the performance target for the delay of one-way video with data rates up to 384 kbps is $<10$ s, which is a large limit we have not tested, but presumably includes any delay from the initial user request until the first required information is received. In Y.1541, delay for video streaming is limited to 1 s. The delay criterion of $<100$ (400) ms for IPTV is the same as for VoIP and does not match the requirements for video traffic in any of the other sources. Reasonable end-to-end delay and jitter values are not considered problematic for IPTV in G.1080, due to STB de-jitter buffers. Nonetheless, it is suggested that jitter for IPTV be kept below 50 ms, both in G.1080 and Y.1541. In compliance with the previous consideration, our tested delays and delay variations did not produce any impact on IPTV quality in our experiments.

Taking into account all the criteria available so far, we can suggest the optimal threshold values. For management and monitoring traffic, packet loss rate should be $<0.007$, and delay $<15$ ms. For VoIP traffic, the required delay value of $<150$ ms is consistent in all of the sources, while we suggest $<0.01$ packet loss rate. For video, IPTV and RTSP streaming traffic require $<10^{-6}$ packet loss rate, but are not limited by reasonable delay, while unicast HLS traffic should be provided with $<0.9$ packet loss rate.

V. CONCLUSION

Based on the known impact of network performance on QoE, when routing traffic, it is possible to select network paths that can ensure the required performance in order to avoid quality degradation of the services delivered. Knowing the criteria for each traffic type is necessary to make an informed routing decision, based on said knowledge. For the purpose of determining the effect of network performance on QoE and for defining routing criteria, we prepared two testbeds. Emulating situations of interest during transmission from source to destination, we tracked how QoE values changed in response to change in network performance metrics. From this data, we were able to extract the values for network performance metrics at which users begin to notice quality degradation.

Based on the network utilization statistics, we extracted relevant service classes, re-examined and redefined their QoS thresholds, and assigned their class priorities. We examined a wide range of traffic types present in a telecommunication operator network, including video traffic with different video delivery techniques, voice traffic, and monitoring/management traffic.

Results from this study can be used for different purposes, e.g., as input for a QoE/QoS analyzer. Based on the network performance measurement results (which are monitored in real time), and on the information on how network performance parameters (delay, jitter, and packet loss) affect individual services, it is possible to perform a QoE/QoS analysis, and gain insight into the quality of the entire network per service.

QoE/QoS analysis provides an insight into which parts of the network are not satisfactory for the transport of individual services. By applying this information to traffic routing, or by eliminating problems that lead to degraded quality, the quality of the service is increased, and thus also the satisfaction of the user. Throughout the paper we referred to the use case of QoE-driven criteria for routing traffic per traffic class in an IP/MPLS network. The current network characteristics of an MPLS tunnel can be compared with the herein obtained threshold values of network parameters, and decisions can be made whether the tunnel is appropriate for routing a particular traffic class. Furthermore, in a real-time measurements system, if a measured network parameter exceeds the obtained threshold value for a traffic type, alarms can be triggered and troubleshooting initiated to detect and resolve the network issue that has arisen.

For future work it is also possible to perform additional extensive real-time measurements and results analysis, which would enable us to verify the proposed decision thresholds, to verify the QoE/QoS models available in the literature, and to propose new appropriate models.

REFERENCES

[1] I. Ramadža, J. Ožegović, and V. Pekić, “Class based tunnel exclusion router architecture,” in 2014 22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2014, pp. 274–278. doi: 10.1109/SoftCOM.2014.7039083
[2] I. Ramadža, J. Ožegović, and V. Pekić, “Network performance monitoring within MPLS traffic engineering enabled networks,” in 2015 23rd International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2015, pp. 315–319. doi: 10.1109/SoftCOM.2015.7341407
[3] Mean opinion score (MOS) terminology, ITU-T Recommendation P.800.1, 2016.
[4] R. C. Streijl, S. Winkler, and D. S. Hands, “Mean Opinion Score (MOS) Revisited: Methods and Applications, Limitations and Alternatives,” Multimed. Syst., vol. 22, no. 2, pp. 213–227, Mar. 2016. doi: https://doi.org/10.1007/s00530-014-0446-1
[5] The E-model: a computational model for use in transmission planning, ITU-T Recommendation G.107, 2015.
[6] Opinion model for video-telephony applications, ITU-T Recommendation G.1070, 2018.
[7] Vocabulary for performance and quality of service, ITU-T Recommendation P.10/G.100, 2017.
[8] K. U. R. Laghari and K. Connelly, “Toward total quality of experience: A QoE model in a communication ecosystem,” IEEE Commun. Mag., vol. 50, no. 4, pp. 58–65, Apr. 2012. doi: 10.1109/MCOM.2012.6178834
Vesna Pekić was born in Split, Croatia, in 1982. She received a M.S. degree in electrical engineering in 2005 and a Ph.D. degree in electrical engineering and information technology in 2013, from the University of Split, Croatia. During her studies she received two Chancellor's and two Faculty awards. After completing her studies, she spent two months at TU Ilmenau in Germany as a part of student practice. Starting in 2005, she worked at the Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, as an external associate for 18 months. In 2007, she joined the faculty as a research assistant. Her research interests include communication protocols, flow control and QoS/QoE on the Internet, digital design, and programmable logic. Dr. Pekić is currently participating in the HGCAL upgrade of the CERN CMS experiment.

Julije Ožegović was born in Split, Croatia, in 1954. He received a M.S. degree in electrical engineering from the University of Zagreb in 1981 and a Ph.D. degree in electrical engineering from the University of Split in 1998. Starting in 1977, he spent two years as a development engineer for digital systems. In 1979, he joined the University of Split's Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, serving as teaching assistant, assistant professor, associate professor, and full professor. His interests include communication protocols, flow control, digital design, programmable logic and ASIC design. Prof. Ožegović is currently participating in the HGCAL upgrade of the CERN CMS experiment. He is the author of two US patents.