DE GRUYTER PIK - Praxis der Informationverarbeitung und -kommunikation

Analysis of Managed and Over-the-Top Streaming Services in Mobile Networks

Journal:	*PIK - Praxis der Informationverarbeitung und -kommunikation*
Manuscript ID:	PIK.2012.0005
Manuscript Type:	Artikel
Date Submitted by the Author:	03-May-2012
Complete List of Authors:	Eisl, Jochen; Nokia Siemens Networks, Lott, Matthias; Nokia Siemens Networks, Varela, Martin; VTT Research Center, Prokkola, Jarmo; VTT Research Center, Kuhn, Gerhard; Nokia Siemens Networks,
Classifications:	
Keywords:	video streaming, quality of experience, measurements
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
photos-authors.zip	

PIK, Vol. 35, pp. 1–7 DOI 10.1515/pik-2012-0025

Analysis of Managed and Over-the-Top Streaming Services in Mobile Networks

Jochen Eisl

Jochen Eisl

Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
jochen.eisl@gmail.com

Gerhard Kuhn

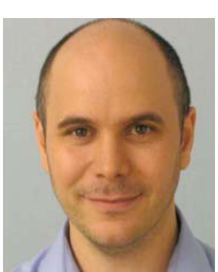
Gerhard Kuhn

Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
gerhard-kuhn@gmx.net

M. Lott

Matthias Lott

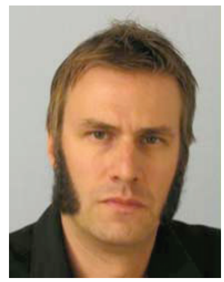
Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
matthias.lott@yahoo.de



Martin Varela

Martin Varela

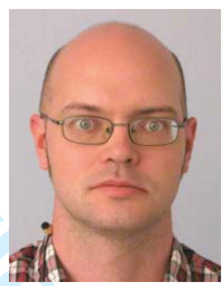
VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
martin.varela@vtt.fi



Jarmo Prokkola

Jarmo Prokkola

VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
jarmo.prokkola@vtt.fi



Toni Mäki

Toni Mäki

VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
toni.maki@vtt.fi

Jukka-Pekka Laulajainen

Jukka-Pekka Laulajainen

VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
jukka-pekka.laulajainen@vtt.fi

Abstract

Optimizing multimedia streaming in 3GPP mobile broadband networks (e.g. HSPA, LTE) in terms of user experience and network resource utilization is one of the most important topics for mobile network operators. The primary reason is that streaming services consume more than half of the total bandwidth. Before designing new concepts and solutions it is fundamental to understand in detail the performance and behavior of multimedia applications for the mobile access. For this reason we have carried out several field experiments to ob-

tain knowledge about typical bandwidth consumption, ability to adapt video bit-rate to changed network conditions, protocol and other service specific issues for the most popular over-the-top and managed video services. We have detected that observed data rates apart from the video encoding and the specific content also depend on client and server implementations. Surprisingly, support for rate adaptation in current implementations of managed and over-the-top streaming services is not very common, in some cases with limited feature set. From the obtained results we conclude that measurements need to continue with a shift from field experiments to lab environment in order to obtain further knowledge of streaming characteristics, which could not be revealed in the current measurements.

1 Motivation & Introduction

Streaming services over IP have attained enormous popularity in the recent years and now contribute significantly to the steady growth of Internet traffic worldwide. The relative growth of traffic in mobile networks is significantly higher due to the ubiquitous usage of mobile devices with excellent display capabilities and built-in cameras. According to a CISCO study the compound annual growth rate for global mobile data traffic between 2011 and 2016 is estimated to be 78% [1]. By end of 2011 streaming video already contributed more than 50% to the total traffic mix of mobile data [1]. Hence the efficient management of streaming video especially within mobile networks is considered to represent a key challenge for mobile network operators. The implications from the technical perspective are manifold. Firstly, a network engineering process might be required for the operator to determine the optimal deployment of network resources based on traffic patterns in the network and the operator business plan for streaming services. Secondly, some additional functions might be needed for optimization of streaming services. This implies the maximization of Quality of Experience (QoE) for the user under possible resource constraints. In conjunction potential bottlenecks in the network have to be detected and possibly the streaming services have to be adapted accordingly. However before detailing solutions for streaming optimization and a network engineering process it is important to understand how streaming services behave in public mobile networks.

For this purpose we have initiated a series of field experiments to analyze streaming protocol details, i.e. implementation aspects and capabilities of a service, such as rate adaptation and other performance characteristics for currently deployed, commercial and non-commercial services. For the described activity both “Over the Top” (OTT) services hosted by a 3rd party provider, and managed services, hosted by the mobile network operator, have been analyzed. The main aspects of both types of services for the measurements are explained in Section II. For the experiment several device and connection types have been used. For each service studied, several measurements were performed, in most cases over different networks, and on different devices, as well as in different ways for each device when appropriate (e.g. browser-based vs. special-purpose “apps”). The details of the measurements setup and constraints are explained in Section III. Then in section IV the results for the different measurements are presented, the most important ones are highlighted explicitly. The focus of the measurements was on observed video bitrates and support for rate adaptation, which might be needed in the case of changing network conditions. Based on these results we conclude the paper with the most important findings in Section V and provide an outlook how to continue the work.

2 Measured Streaming Services Overview

OTT services are currently the most widely spread services in the video domain. These services are provided directly by content providers (and usually over content delivery networks), generally without any arrangement with the network providers sitting between the content and its consumers. Thus, no special treatment is granted to the OTT video streams, and they are provided on a best-effort manner.

Managed services are those in which the content provider and the network provider are the same, or are working in close cooperation. In practice, they mostly refer to services provided by the Internet Service Providers (ISPs) themselves, be it mobile video services provided by mobile carriers, or IPTV services provided by fixed-line ISPs.

The fact that services are managed does not automatically mean that special QoS (Quality of Service) mechanisms are used to provide better performance, however.

In the case of IPTV, it is common for a combination of traffic shaping and DiffServ (Differentiated Services) markings to be used in order to guarantee a certain level of quality to the TV content, while still providing usable network connectivity concurrently[1].

A large variety of services, both OTT and managed, were considered. While most of the measurements were done in Finland, several were done in U.K., Spain and Germany, due to the current lack of managed services provided by Finnish operators. The number of managed streaming services are quite limited up to now (4 out of 18 in our measurements). All of them were tested over 3G/HSPA. For the tested OTT services some include HSPA support. Most of the services were tested with mobile devices. Some of the OTT services did not have support for mobile devices and had to be tested with PC clients only.

3 Measurement Setup

In this section we describe in detail the measurements realized for the current work, including setups, contents, tools and data collection mechanisms.

Two main setups were needed to perform the measurements, given the devices used and their capabilities as well as the tools available. In the case of non-PC, WLAN (Wire-

[1] In some cases, the available bandwidth for network applications is higher when no IPTV services are in use, and the shaping is used only in the case where the IPTV services are actively being used.

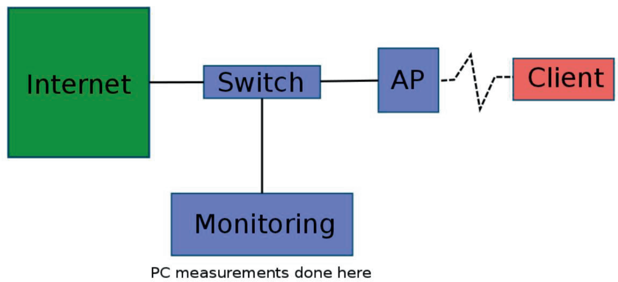


Figure 1 Setup for measurements over WLAN, all platforms.

less Local Area Network)-based measurements, the measurements were done outside the device itself, by mirroring the traffic at a switch before the wireless access point. The access point used was dedicated, and setup in a controlled environment, so it is safe to assume that any performance issues on the video streams were due to elements on the path up to the access point itself. The mirrored traffic was collected and measured on a PC, using Wireshark [2] and Qosmet [3], in a single-point passive measurement configuration. The PC-based measurements were done on the same PC that performed the monitoring both for WLAN and 3G/HSPA (High Speed Packet Access) measurements (see Figure 1). There were, therefore, some setup differences resulting in three final measurement setups. On the one hand, this might introduce some level of difference in the observed performance on different devices. On the other hand, it is usually preferable to perform passive measurements as close to the relevant endpoints as possible, and in this case, those were the devices themselves.

With all these points taken into consideration, we can conclude that the setups were as close to optimal as it was possible to make them.

For the 3G/HSPA measurements, it was only possible to collect the results on the end devices themselves, so the options were limited to those platforms which supported the data collection and measurement tools, namely PC and Android smart phones.

Two main tools were used to capture and measure the studied video streams. Traffic captures on the PC were done with Wireshark [2], which is a proven capture and analysis tool based on pcap [4]. For Android captures, tcpdump and a tool called "Shark" [5] were used. Besides capturing traffic, Wireshark was also heavily used to do protocol-level analysis of what was happening inside the video streams. The second tool used was VTT's Qosmet, which is a passive monitoring tool able to accurately measure several IP-level metrics, as well as some application-level ones. In some cases (especially in some of the U.K.-based measurements), Qosmet measurements could not be done live, so they were performed by replaying sessions captured with Wireshark. To this end, the tcpreplay [6] tool was used on a virtual interface, and the QoS and traffic metrics were taken from there.

Tests over 3G/HSPA were done in Finland, Spain, Germany and the U.K. Within Finland, in addition to normal HSPA access, a special SIM (Subscriber Identity Module) card which is limited to basic UMTS (Universal Mobile Terrestrial Network) rates (i.e. 384 Kbps in the down-link) was used, in order to test for adaptability. For other tests, a router running Linux with the HTB (Hierarchical Token Bucket) queuing discipline was used to limit the link capacity. The goal in both cases was to bring the link capacity below the average bit-rate observed previously for the services, and observe how they would react to this.

The platforms used for the measurements include Symbian, Android2.X, iOS4 for smart phones and Windows XP for PC based measurements. The 3G/HSPA tests were performed both on Android devices and on PCs. In the case of PCs, HSPA dongles were used. Tests performed over WLAN were, for the most part, carried out at VTT's Converging Networks Laboratory with a dedicated WLAN access point so as to avoid interference (either radio or due to other traffic going through the access point).

4 Measurement Results

The main purpose of the measurements was to get insight into streaming service behavior. The two main aspects investigated were consumed video data rate over time and support for adaptation. In addition we wanted to get knowledge about service differentiation, protocol issues, access network aspects and encoding. The following two tables, Table 1 and Table 2 summarize the findings for OTT and managed services.

Subsequently we provide a more detailed explanation of the results and some interpretation, where feasible.

4.1 Protocol , Video Format and Access Network Aspects

Overall, no significant use of DiffServ markings was found in the services measured, both for OTT and managed.

With respect to the video encodings found in the measurements, they were, for the most part, H.264 (Advanced Video Coding – AVC). One service (Livestream) uses MPEG-2, and the YouTube site for Symbian s60 seemingly uses H.263. The Three Free Video service is an outlier, codec-wise, using MP4V (MPEG4-Visual) for video and AMR (Adaptive Multi-Rate) for audio. Otherwise, the audio encoding used the most was AAC (Advanced Audio Coding), often in an MP4-LATM transport, probably to make more efficient use of bandwidth. Other advanced mechanisms for content adaptation, such as Scalable Video Coding (SVC) are not, as of yet, deployed by content providers. This might be attributable to the computational requirements of SVC, which limits the current platforms in which it can perform appropriately.

There is a quite clear-cut distinction between managed and OTT services in terms of the protocols used, with minor exceptions. Managed services use, in all cases, RTSP (Real Time Streaming Protocol) for session management, and RTP/UDP for transport. RTCP (Real Time Control Protocol) sender and receiver reports are also exchanged in all cases where RTP (Real Time Protocol) was used, usually at 1s intervals (which is more frequently than recommended [7]). In

Service Name	Protocol	Encoding	Adaptation
BooxTV (s60 device, live content	RTP/UDP	N/A	Decreases rate and / or does audio only
BooxTV (other devices)	HTTP	H.264 + AAC 320x176	Some (iOS live content, for different bit–rates available
CDOn.com	MMS/HTTP	N/A	No
DailyMotion	HTTP	H.264	No
Facebook Video	HTTP	H.264 + AAC	No
Freebe.tv	RTP/UDP + RTSP	H.264	No
Livestream	HTTP	HTTP H.264 + AAC in MPEG-2TS container	Yes
Vimeo	HTTP	H.264 + AAC	No
Voddler	HTTP	H.264	Only on iOS
YLE Areena (PC)	RTMP/TCP	N/A	No
YLE Areena (other devices)	RTP/UDP + RTSP – HTTP	H.264	Unclear
YouTube (s60)	RTP/UDP + RTSP	RTSP MPEG-4	No
YouTube (other devices)	HTTP	H.264+AAC	No

Table 1 Summary of OTT services studied.

some cases, large number of sender reports containing Source Descriptions (SR SDES) were send as well. One implementation uses application-specific RTCP packets as well. Concerning OTT services there is some variation in the protocols used, but for the most part they use HTTP (Hyper Text Transfer Protocol) as their transport protocol. In particular, HTTP with progressive download (as opposed to actual HTTP streaming) seems to be the norm for the most popular services.

No access-dependent differences in server-side behaviour were noticed during the tests, so the only performance differences when using WLAN vs. mobile accesses are expected to be due to the differences in access networks' performance alone.

Service Name	Protocol	Encoding	Adaptation
Digital+Mobile TV Spain Orange	RTP/UDP + RTSP	H.264 + MP4–LATM	No
Mobile TV Germany T-Mobile	RTP/UDP + RTSP	H.264 + MP4–LATM	Bandwidth ramp up after slow start
Three Free Video U.K	RTP/UDP + RTSP	MP4V + AMR	No
Virgin 3G VoD U.K	RTP/UDP + RTSP	H.264	No

Table 2 Summary of managed services studied.

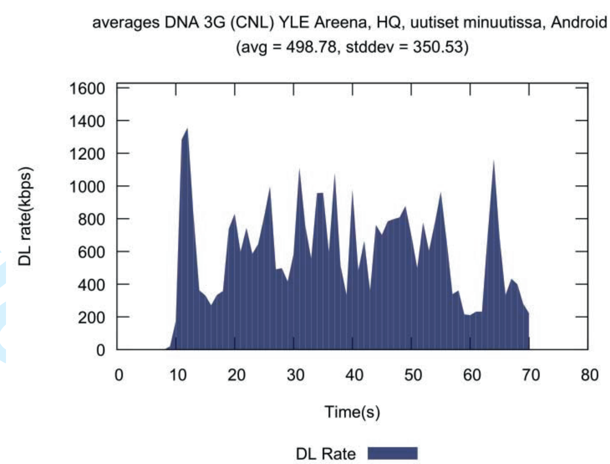


Figure 2 Measured data rate for YLE Arena service.

4.2 Video Services Data Rates

We discovered services with quite variable bit-rates, which cover the whole spectrum of resolutions due to their ability to adapt the quality of the stream. In other cases, such as YouTube, the bit-rates stay more stable throughout the session but are obviously variable with the chosen resolution.

For OTT services using RTP, the transmission rate seems to be quite variable (likely due to variable bit-rate in the stream itself). In Figure 2, the measured down-link rate for YLE Areena with the typical observed bitrate variation can be observed.

BooxTV, on the other hand, maintains a very stable throughput throughout the stream like it can be seen in Figure 3. In the BooxTV case, the actual video bit-rate is variable within the stream, hence the conclusion is that the server is doing some sort of rate control when sending data.

As mentioned above, all managed services tested were RTP-based. Their data rate during play-time was variable itself. For some services (e.g. Virgin's and Three's in the U.K.),

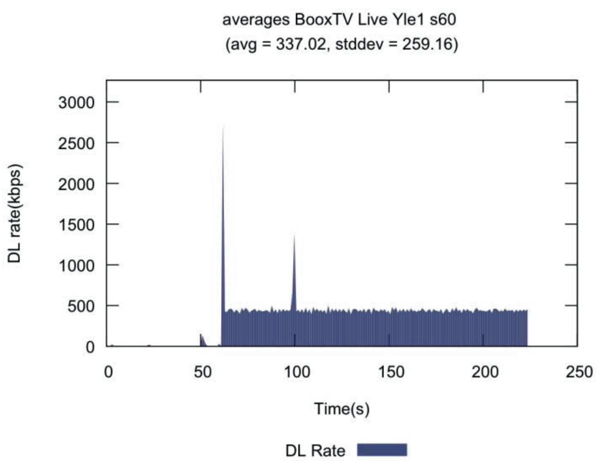


Figure 3 Measured data rate for BooxTV service.

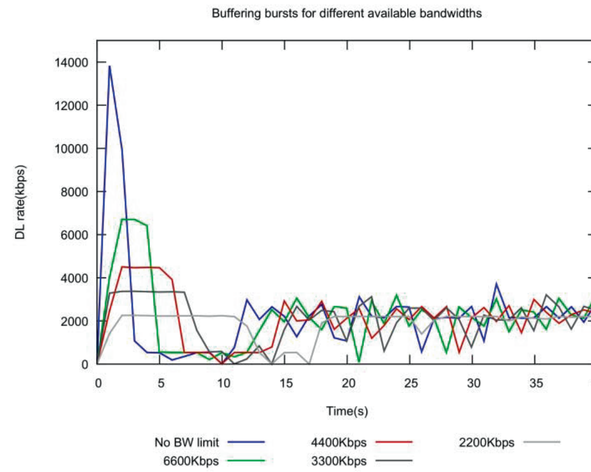


Figure 4 Initial buffering behaviour under several rate limits.

data rates were not so variable as for others (say, Orange's in Spain). Compared to OTT RTP-based services, all the managed services seem to behave in a more regular way concerning bandwidth utilization, and the average throughputs observed correspond quite well to the bit-rates announced in the SDP (Session Description Protocol) data on session initiation.

Those services using HTTP for streaming, often exhibited a rather aggressive initial buffering behaviour, starting with very high-speed initial burst. These bursts could cover between 20–100% of the video file, depending on video resolution (and stream duration) and the service in question (the most common cases were around 25–30%). For example YouTube tends to buffer at least between 8–12 seconds of play-out time before starting playing. The actual times are resolution and device/client dependent, with HD videos buffering in total up to 30 seconds, and lower resolution ones buffering up to 40 seconds. It should be noted that it was not possible to have actual memory usage figures for the buffers, and so the values reported above were measured and/or calculated manually from actual viewing sessions. Figure 4 shows that the buffering is done until the target amount of play-time is reached, as the buffering time becomes longer when the rate is limited.

Another interesting thing to note about HTTP-based services, is that their behaviour varies significantly depending on which client platform is used to access them (e.g. PC vs. iOS vs. Android), and sometimes, even within different versions of the same platform (e.g. Android 2.3 devices do not show such an aggressive buffering behaviour as do those devices running Android 2.2).

4.3 Rate Adaptation

By and large, most services tested did not have any adaptation mechanisms built in. Out of eighteen services tested, only seven showed some type of adaptation mechanism. Of those, two are experimental services with a focus on adaptive streaming. For those that did have adaptation, it was often the case that the adaptation was specific to a certain platform (e.g. Voddler on iOS), or a certain part of the service (e.g. BooxTV only adapts live content). In some cases, the reaction to the decreased bandwidth was just to pause the playback and rebuffer the content as needed. In other cases, the reduction in bandwidth resulted in corrupted streams or services that simply stopped working.

Some services, such as BooxTV and YLE Areena had the "3GPP-Adaptation-Support" field set in the SDP data, but whether or not it was actually used to perform adaptation is unclear. Device-specific adaptations where the server would treat different client devices differently were not observed. If there are any present, they don't seem visible from the client-side.

What follows only applies to the services tested for this study, and as such should not, in principle, be taken to be valid in the general case. That being mentioned, some of the most popular services are covered herein, and therefore the results reported are valid on a very large scale. If not mentioned otherwise the service was tested via WLAN and streaming protocol is based on HTTP (either HLS or MS Smooth Streaming) and iOS devices were used.

Akamai iPhone Video Showcase showed that adaptation is available, but its performance needs improvement. The BooxTV tested for live content enables rate decreases and/or does audio only. The Livestream service supports three levels of adaptation, the smallest one is only audio.

The T-Mobile MobileTV service was tested with Android devices via 3G/HSPA access and is based on RTP/UDP. The interesting observation is that the server ramps up bit-rate after about one minute. Figure 5 shows this example of the streaming, most likely due to the absence of problems in RTCP receiver reports.

The YLE Areena service possibly supports rate adaptation, at least smaller pieces are requested after buffer underrun. The same service was tested over 3G/HSPA with Android device and the service is based on RTP/UDP. For that scenario, audio is prioritized but video becomes unusable when throughput decreases. In this case it is possible that

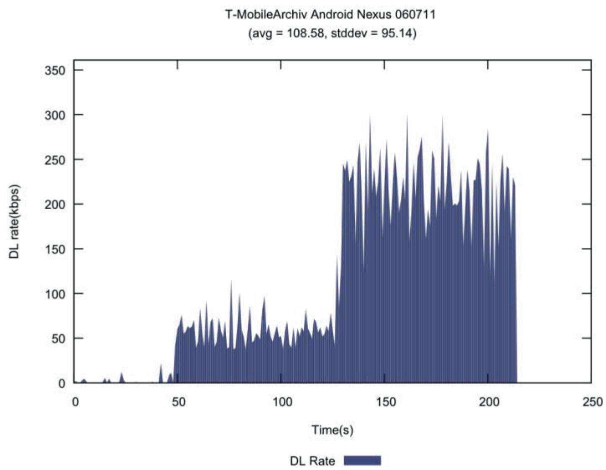


Figure 5 T-Mobile service – Ramp up of video rate.

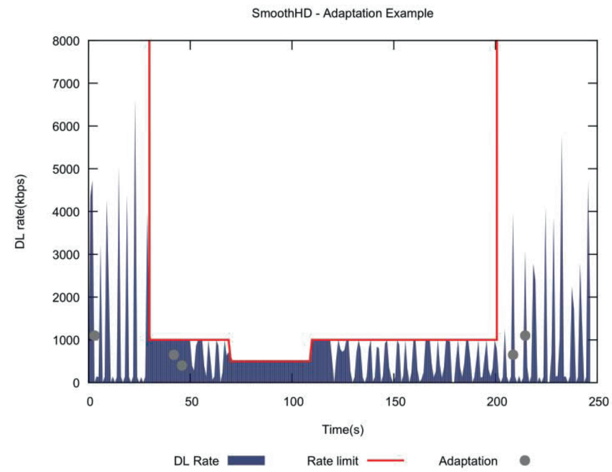


Figure 6 MS SmoothHD – Adaptation Example.

some type of adaptation is being performed by giving priority to the audio stream over the video stream when the bandwidth becomes insufficient.

As mentioned previously, new HTTP-based mechanisms for video streaming have provisions for quality/bit-rate adaptations. Figure 6 shows the adaptation performance of MS SmoothHD. The service tested on PC showed well adaptation performance in terms of quality and responsiveness to changed network conditions. Similarly the Voddler service, which only performs adaptation for iOS clients showed good performance. The application reports four bitrates and adapts accordingly.

It can be seen that the application is quite reactive to even small changes in the available bit-rate. The points marked 'adaptation' in the figure mark the time of an adaptation request (i.e. an HTTP GET command with a different bit-rate requested), and the bit-rate requested. The results obtained show that adaptation mechanisms, save for few exceptions such as Voddler and SmoothHD, still lack both popularity (in terms of their overall use) and technical development in the cases where they are implemented.

5 Conclusions and Next Steps

The purpose of the measurement campaigns reported was to collect information on the current state of video streaming services. This information is not, in many cases, found in the literature, and therefore required intensive labor. The experimental setups were kept as close as possible to real-life conditions, in order to observe actual streaming behaviour.

Our research interest is on optimizing multimedia streaming in 3GPP mobile broadband networks (e.g. HSPA, LTE) in terms of user experience and network resource utilization. Before designing new concepts and solutions it is fundamental to understand in detail the performance and behavior of multimedia applications for the mobile access. Our results provide insight into typical bandwidth consumption, ability to adapt video bit-rate to changed network conditions, protocol and other service specific issues for the most popular over-the-top and managed video services. We have used several mobile devices with different OS platforms and PC as well. Due to sparse offering in managed video services we conducted experiments in four different European countries. Performing experiments in Finland, Spain, Germany and the U.K. enhances the validity of the results obtained and provides a European-level view of the current state of video streaming services.

While the use of adaptation mechanisms is sparse, it would be surprising if they weren't to become more commonplace in the coming years, as they have the potential to deliver a significantly better experience to end users via graceful degradation. Like mentioned only a few services do implement some sort of adaptation, with varying degrees of success, performance-wise. In most cases, a sudden restriction in bandwidth will result in either pausing and re-buffering, or just breaking the stream. Some services do try to adapt, using different mechanisms.

In any case, the number of OTT services clearly outnumbers that of managed services as of this writing. In terms of performance (i.e. quality), the subjective opinions of the people doing the measurements was that the performance for managed services was not significantly better than for OTT services. From a technical perspective, managed services would be simpler to provide with suitable quality guarantees. For OTT services quality is variable and oftentimes the same content is available at several different resolutions from the same content provider, so clients can choose the most suitable one for their context (in most cases, this choice needs to be made explicit). For some platforms, dedicated applications exist for a wide variety of services, and some services are implemented differently for different platforms. Unlike the case of managed services, where SDP data was readily available, extracting media information for OTT services was not always feasible, e.g. due to encryption of data.

The results we obtained indicate that support for streaming optimizations are not widely offered or with limited feature sets only by today's common streaming services and networks. So far we obtained a look from the user perspective by collecting measurement data at the terminal and obtaining user experience for rate adaptation support. Based on these results we decided to continue measurement series in mobile network lab environment to extract information of streaming server behavior in conjunction with the information exchanged with the user terminal.

References

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016, http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html.

[2] Various authors. The Wireshark Protocol Analyzer. http://www.wireshark.org, 2011.

[3] J. Prokkola. Qosmet. http://www.cnl.fi/qosmet.html.

[4] Various authors. Libpcap: A Portable C/C++ Library for Traffic Capture. http://www.tcpdump.org, 2011.

[5] E. Kuštans. Shark for Root – Traffic Sniffer for Android. https://market.android.com/details?id=lv.n3o.shark, 2011.

[6] A. Turner. Tcpreplay. http://tcpreplay.synfin.net/, 2011.

[7] H. Schulzrinne, et al. "RFC 3550 – RTP: A Transport Protocol for Real-Time Applications". http://www.ietf.org/rfc/rfc3550.txt, 2003.

J. Eisl
Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
jochen.eisl@gmail.com

G. Kuhn
Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
gerhard-kuhn@gmx.net

M. Lott
Nokia Siemens Networks
St.-Martin-Str. 76
81541 Munich
Germany
matthias.lott@yahoo.de

M. Varela
VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
martin.varela@vtt.fi

J. Prokkola
VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
jarmo.prokkola@vtt.fi

T. Mäki
VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
toni.maki@vtt.fi

J.-P. Laulajainen
VTT Technical Research Centre of Finland
Kaitoväylä 1
90570 Oulu
Finland
jukka-pekka.laulajainen@vtt.fi