Transcoding and Streaming-as-a-Service for improved Video Quality on the Web

Christian Timmerer
Alpen-Adria-Universität Klagenfurt / bitmovin GmbH
Universitätsstraße 65-67
9020 Klagenfurt, Austria
christian.timmerer@itec.aau.at / christian.timmerer@bitmovin.com

Daniel Weinberger
bitmovin GmbH
Lakeside B01
9020 Klagenfurt, Austria
daniel.weinberger@bitmovin.com

Martin Smole
bitmovin GmbH
Lakeside B01
9020 Klagenfurt, Austria
martin.smole@bitmovin.com

Reinhard Grandl
bitmovin GmbH
Lakeside B01
9020 Klagenfurt, Austria
reinhard.grandl@bitmovin.com

Christopher Müller
bitmovin GmbH
Lakeside B01
9020 Klagenfurt, Austria
christopher.mueller@bitmovin.com

Stefan Lederer
bitmovin GmbH
Lakeside B01
9020 Klagenfurt, Austria
stefan.lederer@bitmovin.com

ABSTRACT
Real-time entertainment application such as the streaming of audio and video is responsible for the majority of today's Internet traffic. The transport thereof is accomplished over the top of the existing infrastructure and with MPEG-DASH interoperability is achieved. However, standards like MPEG-DASH only provide the format definition and the actual behavior of the corresponding implementations is left open for (industry) competition. In this demo paper we present our cloud encoding service and HTML5 adaptive streaming player enabling highest quality (i.e., no stalls and maximum media quality under the given conditions) and low start-up delay (for both live and on demand) at low (networking) costs. The demo comprises a feature-rich set of DASH-based application scenarios including much faster than real-time encoding/transcoding in the public cloud, live (device-to-device) streaming, support for digital rights management and advertisements, ultra high-definition and high frame rates, virtual reality, and 360° videos.

CCS Concepts
• **Information systems→Information systems applications→** Multimedia information systems→Multimedia streaming

Keywords
Transcoding and Streaming-as-a-Service, Adaptive Media Streaming, Quality of Experience, bitmovin, MPEG-DASH.

1. INTRODUCTION
The amount of the global Internet traffic has considerably increased over the past five years [1] and the streaming of audio and video currently – in some cases – accounts for more than 70% of the Internet traffic [2] resulting in a complex set of requirements for the various players within multimedia content delivery ecosystems. Interestingly, the trend analysis in [1] shows that by 2019 traffic originating from wireless/mobile devices will exceed the traffic from wired/stationary devices.

The goal to make video a "first class citizen" on the Internet (and consequently the Web) has become reality thanks to improved network capacities and capabilities, more efficient (video) codecs (e.g., AVC, HEVC), powerful end user devices (e.g., computers, laptops, smart devices (TVs, tablet, phones, watches)), and corresponding Web technologies (e.g., HTML5 and associated recommendations). Additionally, the Hypertext Transfer Protocol (HTTP) – "the narrow waist of the future internet" [3] – is more and more superseding the Internet Protocol (IP) as a primary mean for transportation (although and of course IP is still in use) but the focus is shifting more towards applications and services rather than transport/network layer optimizations. In this context, the video streaming over HTTP has emerged in the last few years and opened a punch of new research and engineering opportunities for both industry and academia which led to the standardization of MPEG-DASH [4] allowing for interoperability among different vendors and flexibility to develop new services on top of it.

In this demo paper we briefly highlight the scientific and engineering concepts behind our cloud encoding service and HTML5 adaptive streaming player enabling highest quality (i.e., no stalls and maximum media quality under the given conditions) and low start-up delay (for both live and on demand) at low (networking) costs. We then show the novelty of this work based on a plethora of use cases supported by our technology and provide an overview of what will be demonstrated at the conference. Finally, we provide a conclusion and outlook.

2. BACKGROUND
In this section we highlight the scientific and engineering concepts behind our cloud encoding service and HTML5 adaptive streaming player.

Hardware encoder manufacturers have adopted their product lines to support the dynamic adaptive streaming over HTTP but suffer from the inflexibility to provide scalability on demand, specifically for event-based live services that are only offered for a limited period of time. The cloud computing paradigm allows for this kind of flexibility and provide the necessary elasticity in

MMSys'16, May 10-13, 2016, Klagenfurt, Austria

ACM 978-1-4503-4297-1/16/05.

DOI: http://dx.doi.org/10.1145/2910017.2910637

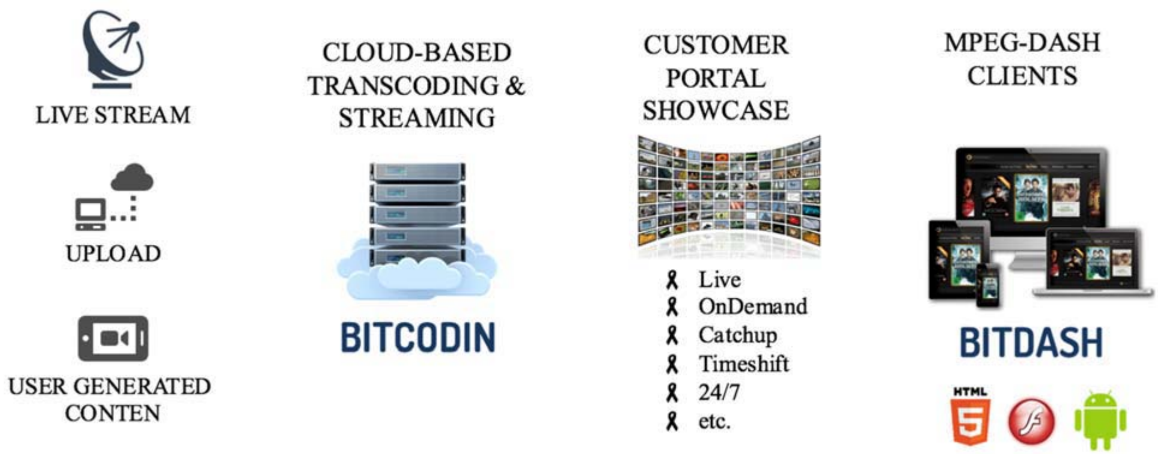

Figure 1: High-Level System Architecture for Live Transcoding and Streaming-as-a-Service.

order to easily scale with the demand required for such use case scenarios. Additionally, service providers like Netflix and YouTube have demonstrated the use of cloud encoding services also for on demand use cases.

The high-level system architecture is depicted in Figure 1 which takes the live (or on demand) multimedia content as an input and transcodes it to multiple content representations in real-time on standard infrastructure-as-a-service (IaaS) cloud environments according to the requirements of the customer in terms of codecs, resolutions, bitrates, quality, etc. These requirements are expressed through an application programming interface (API) exposed to the customer. The resulting manifest describing the individual content representations and primary input for the adaptive streaming client is incorporated within the customer's Web portal offering the service to the actual clients (i.e., end users). The transport within the network is conducted utilizing standard Content Delivery Network (CDN) infrastructures. The heterogeneous clients request the multimedia segments based on the manifest received prior to the streaming and adapt themselves to the context conditions such as fluctuating network bandwidth.

The MPEG-DASH standard defines the media presentation description (MPD) as well as segment formats and deliberately excludes the specification of the client behavior, i.e., the implementation of the adaptation logic, which determines the scheduling of the segment requests, is left open for competition. In the past, various implementations of the adaptation logic have been proposed [5] both within the research community and industry deployments/products. In any case, the behavior of the adaptation logic directly impacts the Quality of Experience (QoE) which – in case of DASH – mainly depends on the initial/start-up delay, buffer underruns also known as stalls, quality switches, and media throughput [5]. The goal is certainly to enable streaming at highest quality (e.g., highest audio-visual quality and no stalls) with low start-up delay (for both live and on demand), and at low (networking) costs.

Please note that our approach explicitly targets MPEG-DASH as the primary multimedia format but in practice, however, there is also a need to support Apple's HTTP Live Streaming (HLS) for iOS-based devices. In addition to DASH and HLS we adopt the following main technologies for our streaming service. HTML5 for video playback in combination with Media Source Extensions (MSE) and Encrypted Media Extensions (EME). The client implementation relies on JavaScript/ECMAScript which makes plugins (like Flash) unnecessary.

The novelty of our approach can be described by our cloud encoding services which exploits a public cloud environment that provides the elasticity and scalability to offer high quality streaming services. It offers a rich REST API for various integration and configuration purposes such as various inputs, outputs, encoding profiles, jobs, manifests, thumbnails, statistics, payment, live streaming, and numerous API clients for Java, JavaScript/ECMAScript, Ruby, Python, PHP, and NodeJS. On the playback side also a rich API is offered allowing for player setup, playback commands, player status, metrics, metadata handling, skinning, Chromecast support, and event handlers. Finally, we support a plethora of use cases described in the next section.

3. DEMO DESCRIPTION

In this section we give a brief overview of the demos to be provided at the conference.

3.1 Faster than Real-Time Encoding

The encoding mechanism utilizes the flexibility and elasticity of existing cloud infrastructure-as-a-service (IaaS) providing scalability on demand when it is needed. In particular, cloud instances are requested and utilized depending on the demand in order to satisfy real-time requirements and even beyond, i.e., transcoding to various content representations ranging from standard definition resolution to ultra high definition resolution multiple times faster than real-time. A screen shot is shown here (with 141x real-time transcoding), which reveals the performance of being much faster than real-time and a preview of the results while the transcoding is still in progress.

3.2 Live Streaming

Everything in our workflow is designed for the live use case which enables the streaming of both professional and user-generated content to a plethora of heterogeneous devices with digital video recording (DVR) capabilities. In particular, we support (and have deployed) 24/7 live services and event-based live services. The latter fully exploits the cloud computing paradigm as the cloud environment is only used for a given amount of time which drastically reduces costs compared to fixed hardware installations.

The content is provided using the proprietary Real-Time Messaging Protocol (RTMP) push, which is a de-facto standard used within the industry to push live content over the Internet. Other open standards such as HTTP/2 push could be a replacement but it is not yet widely available, as it has been only standardized recently [6]. Finally, we will demonstrate how easy it is to setup a live streaming session with mobile devices as a source which effectively demonstrates a device-to-device adaptive streaming over HTTP using our services and tools as the underlying platform.

3.3 Support for DRM and Advertisement

Another demo shows the support for Digital Rights Management (DRM) via MPEG Common Encryption (CENC) [ISO/IEC 23001-7:2016] in HTML5, which is enabled by the Encrypted Media Extensions (EME) of modern browsers, like Chrome, Firefox, Internet Explorer, etc., and also used in Flash for other browsers or older versions. The demo will also show the support for the proprietary DRM schemes such as PlayReady, Access/Primetime, FairPlay and Widevine.

In order to monetize content the player can be configured with pre-, mid-, and post-roll ads according to the Digital Video Ad Serving Template (VAST) 3.0 [7] which will be part of our demonstration.

Finally, we will show means for skinning the Web player according to customer's individual design and usability needs.

3.4 UHD Streaming and High Frame Rates

An important aspect for future deployments will be the support for Ultra-High-Definition (UHD) and high frame rates (HFR). Thus, we will demonstrate the playback of 4K and 8K resolution videos as well as full HD videos with 60 frames per second using the state-of-the-art Advanced Video Coding (AVC) standard. For resolutions beyond full HD the new High Efficiency Video Coding (HEVC) standard [8] has been developed that promises half the bitrate at the same visual quality compared to AVC. In this context we will show a full end-to-end transcoding and streaming solution using HEVC with the aim to support future UHD service at high quality but also using it for lower resolution in order to further decrease the bitrate and, thus, reducing (networking) costs. At the time of writing this paper, our API supports an HEVC encoding of up to 7680x4320 pixels and 120 frames per second which is truly UHD (actually our solution supports higher resolutions and framerates but this is not yet supported on any consumer electronic devices).

3.5 Virtual Reality and 360° Videos

Finally, we will demonstrate new uses in the area of virtual reality (VR) and 360° videos as capturing this type of content is becoming easier thanks to the availability of commodity hardware. Therefore, we have developed a plugin, which enables our adaptive streaming player solution to playback VR and 360° video natively in HTML5 to enable an immersive experience for everyone. To make use of the VR and 360° video support in various usage scenarios, the player offers different rendering/projection modes which fit for desktop and mobile devices as well as for the usage in combination with VR headsets.

4. CONCLUSIONS

In this demo paper we have shown the potential of HTTP streaming and the MPEG-DASH standard in various usage scenarios. Our implementation efforts within public cloud infrastructure provides the necessary elasticity and scalability in order accommodate today's and future customer needs. The HTML5-based client framework allows an integration in all modern browser platforms without the need for a plugin which was previously a major burden (e.g., Flash was never supported on iOS).

We have also shown that the deployment of ultra high-definition services or emerging application scenarios such as virtual reality and 360° videos are basically supported by the underlying technology infrastructure. The optimization towards Quality of Experience and network utilization is declared as future work.

5. ACKNOWLEDGMENTS

This work was supported in part by the Austrian Research Promotion Agency (FFG) under the AdvUHD-DASH project (no. 848132) and the EU-FP7-ICoSOLE (no. 610370).

6. REFERENCES

[1] Cisco. Cisco Visual Networking Index: Forecast and Methodology, 2014—2019. Technical report, Cisco, 2015.

[2] Sandvine. Global Internet Phenomena Report Africa, Middle East, and North America. Technical report, Sandvine Intelligent Broadband Networks, 2015.

[3] Popa, L., Ghodsi, A., and Stoica, I. 2010. HTTP as the narrow waist of the future internet. In Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks (Hotnets-IX). ACM, New York, NY, USA, Article 6 , 6 pages. DOI=http://dx.doi.org/10.1145/1868447.1868453

[4] Sodogar, I., The MPEG-DASH Standard for Multimedia Streaming over the Internet, IEEE Multimedia*,* Vol. 18, No. 4, Oct.-Dec. 2011, pp. 62-67.

[5] Timmerer, C., Weinberger, D., Mueller, C., Lederer, S. Ultra-High-Definition-Quality of Experience with MPEG-DASH, Broadcast Engineering Conference (BEC), NAB2015, Las Vegas, NV, USA, April 2015.

[6] Belshe, M., Peon, R., Thomson, M., (eds.), Hypertext Transfer Protocol version 2, February 2015.

[7] IAB. Digital Video Ad Serving Template (VAST) 3.0. July 2012. http://www.iab.com/guidelines/digital-video-ad-serving-template-vast-3-0/ (last access: February 2016)

[8] Sullivan, G., Ohm, J.-R., Han, W.-J., and Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. Circuits and Systems for Video Technology, IEEE Transactions on, 22(12):1649–1668, December 2012.