Adaptive Bit Rate (ABR) Multistreaming: HEVC Encoding and MPEG-DASH Content Delivery Networks

Greg Mirsky PhD, VP Product Development
Nick Terterov PhD, Chief Scientist
Sam Blinstein PhD, Member of Technical Staff

April 8, 2015

Vanguard Video, LLC.
974 Commercial Street, Suite 200
Palo Alto, CA  94303
(650) 961-3098 (voice)
(650) 292-2340 (fax)
ask@vanguardvideo.com
http://www.vanguardvideo.com
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>SECTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>3</td>
</tr>
<tr>
<td>H.265 and MPEG-DASH</td>
<td>3</td>
</tr>
<tr>
<td>V.265 MULTISTREAM API</td>
<td>4</td>
</tr>
<tr>
<td>COMPUTATIONAL RESOURCES OPTIMIZATION</td>
<td>4</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>7</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>7</td>
</tr>
</tbody>
</table>
INTRODUCTION

The purpose of this whitepaper is to provide a description and analysis of HEVC-based video adaptive bit rate (ABR) multistreaming with a focus on integration into Content Distribution Networks (CDNs), and in particular ones based on MPEG-DASH. For the past few years, live and offline video streaming to multiple devices has been an ongoing challenge. The lack of network capacity, infrastructure limitations, and bandwidth instabilities contributed to problems with synchronized and simultaneous streaming to multiple devices. Our focus will be on providing solutions to this challenge using a combination of existing industry standards and Vanguard Video’s latest HEVC codec technology.

H.265 AND MPEG-DASH

Recent developments in multimedia technology have provided two major tools to solve the problem of optimizing the delivery of synchronized video streams across multiple devices: 1) the next generation video coding standard HEVC to reduce bandwidth, and 2) MPEG-DASH to manage the streaming of synchronized resolutions and bitrates across varying network bandwidths to multiple devices.

The latest video coding standard HEVC offers significant compression efficiency gains over its predecessor AVC, which can be used to support higher quality video delivery over limited bandwidth networks, especially for mobile devices. The increase in compression efficiency comes primarily from the large quantity of new encoder tools including many more fine-grain modes: video block partitions, many more prediction modes, and 32x32 transforms to name a few. Additionally, both software and hardware infrastructure support for HEVC is now available, and, in the last couple of years, multiple broadcast companies have demonstrated HEVC delivery over professional broadcasting networks.

MPEG-DASH was standardized before HEVC and is already widely used for the distribution of AVC content. MPEG-DASH works by splitting content into a collection of file segments, each containing a short portion of content. A corresponding Media Presentation Description (MPD) file contains timing, segment URL, and media properties such as resolution, framerate, bitrate, and codec profiles. The segments and the MPD are stored on an HTTP server and streamed to MPEG-DASH clients connecting to the server. As the content is streamed to an MPEG-DASH client, the client can dynamically and adaptively select segments of different resolutions and bitrates depending on network conditions and device capabilities. Although MPEG-DASH is codec and media format agnostic, the MPEG-DASH standard defines guidelines for implementing interoperable adaptive streaming services. In particular, it describes specific media formats for use with the ISO Base Media File Format (e.g., MP4) or MPEG-2 Transport Stream containers. These containers already support the inclusion of HEVC elementary video streams, so integrating HEVC into an MPEG-DASH workflow is possible within the existing standards.

MPEG-DASH targets OTT delivery and CDNs, but it is also used in some broadcast environments as a replacement for MPEG-2 TS-based workflows. Through the exhaustive descriptions available in the MPD, MPEG-DASH clients can determine which media segments best fit their user’s preferences, device capability, and network conditions, guaranteeing both a high quality viewing experience and support for the next-generation of video services. Moreover, there is no constraint on the duration of file segments. For these reasons, there is an industry trend to use MPEG-DASH in CE devices such as smart televisions, gaming consoles, smartphones and tablets.
V.265 MULTISTREAM API

Very early on in the development of HEVC, Vanguard Video realized the need for true adaptive bitrate (ABR) multistreaming support in HEVC as a tool for content preparation for multistreaming services. In Version 3.0 of V.265, Vanguard Video’s software HEVC SDK solution, we introduced several API extensions supporting multistream encoding. This architecture allows for the encoding of a single master source video into multiple, GOP-aligned streams of various resolutions and bitrates with a single encoder instance. Moreover, newly exposed encoder input settings allow for the specification of individual settings and flags for each of the streams. This functionality offers two significant advantages for developing a multistreaming workflow. First, the architecture guarantees that the multiple streams are GOP-aligned, an essential requirement for any multistreaming workflow, and in particular any based on MPEG-DASH. Secondly, it simplifies the encoding process to a single input source and encoder instance, reducing command-and-control and resource management. Although GOP-alignment is possible to achieve using multiple encoders, one must often sacrifice certain critical functionality whose variance across the separately encoded streams cannot guarantee alignment, for example, scene-change detection. With V.265 these features can be used to produce multiple streams while ensuring GOP-alignment with all the HEVC encoder features, including scene change detection synchronization across streams. Additionally, V.265 multistreaming support optimizes the allocation and management of various internal operations such as video resizing, coarse motion estimation, and GOP-alignment and avoids duplication of common settings and operations.

In short, the V.265 multistreaming architecture offers a robust and maximally efficient implementation of HEVC-based stream encoding for multistreaming delivery. Moreover, because the encoder produces elementary HEVC video streams, it is agnostic with respect to the specific multistreaming delivery method. By using commercial tools or open source tools such as MP4Box to segment the video and audio streams and mux them into containers, V.265 is easy to integrate into a variety of streaming deployments such as MPEG-DASH, Apple HTTP Live Streaming (HLS), Microsoft Smooth Streaming, or Adobe HTTP Dynamic Streaming.

See Figure 1 below for a diagram of a hypothetical MPEG-DASH-based multistreaming encoding workflow integrated with the V.265 encoder.

COMPUTATIONAL RESOURCES OPTIMIZATION

Conceptually, one of the potential advantages of a multistreaming encoding solution is computational savings from sharing specific components of the processing pipeline. Unfortunately, this is not the case in practice because the computationally significant components are strongly resolution or stream dependent. The main areas include motion estimation, encoding decision-making, and entropy encoding. Although coarse motion estimation occurs on a reduced-resolution version of a given source video, computationally-expensive refined motion estimation is directly dependent on resolution. Therefore, results of refined motion estimation cannot automatically be applied to lower resolution versions of the master source. Whereas motion vectors determined from coarse motion-estimation on reduced-resolution frames can be scaled as seeds for the refined motion-estimation engine, the inaccuracy of this process is precisely why refined motion-estimation on the true video resolution must
be performed. In light of this, some savings can be obtained when one resolution is a multiple of another (e.g., 3840x2160 and 1920x1080) and indeed this is done within the V.265 multistreaming framework.

As mentioned above, the strong gains in compression efficiency offered by HEVC occur in the increased variety of encoding decisions that the encoder makes to optimize bit allocation. But this comes at a computational cost which unfortunately cannot be significantly reduced across different streams: video partitions, intra-modes, and transform trees are all directly dependent on video resolutions and bitrates. Similarly, entropy encoding is specific to the individual bit makeup of the encoded stream so the context-adaptive arithmetic coding of the CABAC engine cannot simply be copied for different streams. Even minor bit-differences due to multithreading execution in streams encoded from the exact same source could produce different results after entropy encoding.

All-in-all, expected computational savings from an HEVC-based multistreaming encoding framework amount to about 5-10% in CPU load reduction, as compared to using separate encoders for each stream. Despite the modest computational performance savings available from multistreaming, Vanguard Video’s multistreaming implementation yields optimally synchronized streams that are nearly impossible to generate using separate encoders.
Adaptive Bit Rate (ABR) Multistreaming: HEVC Encoding and MPEG-DASH Content Delivery Networks

Figure 1: Hypothetical MPEG-DASH multistreaming workflow with V.265 HEVC encoding.
CONCLUSIONS

V.265’s multistreaming support offers a robust and simplified encoding workflow for integration with real-world multistreaming delivery methods such as those based on MPEG DASH. Supporting multiple streams of varying resolutions, bitrates, and settings from a single source in a single encoder instance, which guarantees stream GOP-alignment and offers computational savings across shared processes, V.265 offers providers the opportunity to combine the advancements of HEVC coding with the versatility of MPEG-DASH to offer true next-generation video delivery across the multitude of networks and devices in use today.

REFERENCES


About Vanguard Video

Vanguard Video is a supplier of professional, broadcast quality H.265/HEVC and H.264 codec technologies to top tier customers around the world. With deep codec expertise, unparalleled performance/quality, and world class support and integration services, Vanguard Video has helped its customers capitalize on many first to market opportunities by pioneering advanced compression technologies including the release of the world’s first commercially deployed H.265/HEVC service. Vanguard Video technologies support a wide range of platforms including x86, ARM, and OpenCL acceleration for GPUs. [www.vanguardvideo.com](http://www.vanguardvideo.com)

Corporate Headquarters
974 Commercial Street, Suite 200
Palo Alto, CA, 94303 USA
Phone: +1 (650) 961-3098
Fax: +1 (650) 230-4904