Future DASH Applications: a Survey

Koffka Khan
Department of Computing and Information Technology the University of the West Indies, Trinidad and Tobago, W.I
Email: koffka.khan@gmail.com

Wayne Goodridge
Department of Computing and Information Technology The University of the West Indies, Trinidad and Tobago, W.I
Email: wayne.goodridge@sta.uwi.edu.com

Abstract

In recent years DASH has become the de facto standard for video streaming. However, applications continue to demand more bandwidth in an effort to satisfy users. This is based on applications such as 360° and immersive Virtual Reality/Augmented Reality (VR/AR) video. This paper explores current state of the art and introduces the reader to some of these DASH-aware bandwidth intensive applications.

Keywords - DASH; applications; bandwidth; video; streaming; 360°; VR; AR.

I. INTRODUCTION

The concept of adaptive video streaming is based on the idea to adapt the bandwidth required by the video stream to the throughput available on the network path from the stream source to the client [13]. These algorithms can live at the server [1], at an intermediate network device [15] or at the client [9]. With the client-side approach it is the player that decides what bitrate to request for any fragment, improving server-side scalability [1]. A benefit of this approach is that the player can control its playback buffer size by dynamically adjusting the rate at which new fragments are requested. The adaptation is performed by varying the quality of the segments. Multiple video segments constitute a video stream lasting from as little as 2 seconds to as much as 10 seconds of chunk delivery rate. Segments are encoded and stored on the server in numerous quality versions, termed representations. Each version has a unique resolution, bitrate and/or quality. A client downloads segments using HTTP GET statements [12]. However, with adaptive streaming a client might request subsequent segments at different quality levels to manage varying network conditions, based on an estimation of bandwidth. To do this it uses a manifest file that contains information about the video segments.

Protocols and standards such as MPEG Dynamic Adaptive Streaming over HTTP (DASH), Apple HTTP Live Streaming (HLS), Microsoft Smooth Streaming (MSS) or Adobe HTTP Dynamic Streaming (HDS) typically use a media playlist that contains a list of uniform resource identifiers (URIs) that are addresses to media segments [21]. The process of determining the ideal representation for each segment to enhance the user’s experience is pivotal to adaptive streaming. The controller algorithm estimates the network bandwidth and chooses the next bitrate level corresponding to the available network bandwidth. Variations in the available bandwidth will result in jerky playback and disruption of the video playback if the throughput falls below the bit rate requirement of the video. This is the major challenge in adaptive video streaming. Selecting appropriate bitrate levels helps to maximize the user experience. Generally, higher bitrates and resolutions will give better user experience. For example, if a client approximates that there is 9.5Mb/s available in the network, it might request the server to stream video compressed to the highest video rate available, 9.5Mb/s, or the next rate below, 9.3Mb/s. If the client picks a video rate that is too high, the viewer will experience annoying re-buffering events; if they pick a streaming rate that is too low, the viewer will experience poor video quality. In both cases, the experience degrades [30], [10], [22], and user may take their viewing elsewhere [5]. It is therefore important for a video streaming service to select the highest safe video rate. [8]

Adaptive streaming uses the HTTP/TCP protocol stack to transmit video Web traffic. Thus, the development of this wave of HTTP-based streaming applications is not referred to as adaptive streaming over HTTP. The use of HTTP/TCP protocols for video streaming is because of the advantages that HTTP/TCP offers. It allows standard web servers and caches to be used increasing its’ cost effectiveness. Another advantage is that all firewalls are configured to support HTTP connections [13]. In addition, is allows better scaling as HTTP is stateless and the streaming session is managed by the client, thus reducing the load on the server. However, HTTP/TCP use reveals further challenges as adaptation is on top of TCPs congestion control algorithm, which forms nested control loops. As the throughput of the TCP connection depends on both the link capacity and the amount of congestion, the throughput can vary significantly over time [18]. Video over IP is becoming more and more important as we move further into the twenty-first century. The Internet is still growing rapidly, and more uses are being found for video users. These include real-time online visual assistance, video learning, live event streaming, smart HDTVs, mobile phones, gaming devices, computers and visual communication among others. As the content quality is improving to meet end-user demands the bandwidth requirement for such devices is rapidly increasing. This paper aims to introduce and categorize some of the main applications that utilize DASH over the Internet.

This work consists of four sections. Section II presents a categorization of present day DASH applications. Section III gives examples of these applications. Finally, the conclusion is given in Section IV.
II. DASH APPLICATIONS

DASH-based future applications are categorized into 360°, Volumetric, Virtual Reality (VR), Imaging, Immersive, Zoomable and Web 3D. Further, Volumetric is broken down into Augmented Reality (AR) 3D and VR. VR is broken down into VR 3D, VR 2D and 360° VR. Immersive is broken down into Immersive 3D and Virtual Transcendence Experience (VTE). Finally, Cloud 3D is a component of Web 3D.

Fig. 1. Future DASH applications.

III. EXAMPLES OF DASH APPLICATIONS

This review of SAND-based approaches includes those inspired by SAND and actual implementations. It should be noted that SAND is an adaptive video streaming architecture so that both inspired and direct implementations may or may not include all the features included in the architecture.

A. Immersive

1. Immersive 2D

Immersive video offers the freedom to navigate inside virtualized environment. Instead of streaming the bulky immersive videos entirely, a viewport or field of view (FoV) adaptive streaming is preferred (cf. Figure 2). We often stream the high-quality content within current viewport, while reducing the quality of representation elsewhere to save the network bandwidth consumption. We then could refine the quality when focusing to a new FoV. Therefore, in their work, authors have attempted to model the perceptual response of the quality variations (through adapting the quantization and spatial resolution) with respect to the refinement duration and reach at a product of two closed-form exponential functions that well explain the joint quantization and resolution induced quality impact. Analytical model is also cross-validated using another set of data with both Pearson and Spearman's rank correlations over 0.98. Authors work would be devised to guide the bandwidth-quality optimized immersive video streaming.

Fig. 2. Delivery system architecture [19].

Immersive video offers the freedom to navigate inside the virtualized environment. Instead of streaming the entire bulky content, a viewport or field of view (FoV) adaptive streaming is preferred. We often stream the high-quality content within current viewport, but degraded-quality representation elsewhere, so as to reduce the network bandwidth consumption. We then could refine the quality when focusing to a new FoV. Therefore, in their work, authors have attempted to model the perceptual response of the quality variations (through adapting the quantization and spatial resolution) with respect to the refinement duration and reach at a product of two closed-form exponential functions that well explain the joint quantization and resolution induced quality impact. Analytical model is also cross-validated using another set of data with both Pearson and Spearman's rank correlations over 0.98. Authors work would be devised to guide the bandwidth-quality optimized immersive video streaming.

2. Immersive 3D

High-quality point clouds have recently gained interest as an emerging form of representing immersive 3D graphics. Unfortunately, these 3D media are bulky and severely bandwidth intensive, which makes it difficult for streaming to resource-limited and mobile devices. This has called researchers to propose efficient and adaptive approaches for streaming of high-quality point clouds.

In their paper, authors run a pilot study towards dynamic adaptive point cloud streaming and extend the concept of dynamic adaptive streaming over HTTP (DASH) towards DASH-PC, a dynamic adaptive bandwidth-efficient and view-aware point cloud streaming system. DASH-PC can tackle the huge bandwidth demands of dense point cloud streaming while at the same time can semantically link to human visual acuity to maintain high visual quality when needed (cf. Figure 3). In order to describe the various quality representations, authors propose multiple thinning approaches to spatially sub-sample point clouds in the 3D space and design a DASH Media Presentation Description manifest specific for point cloud streaming. Authors initial evaluations show that they can achieve significant bandwidth and performance improvement on dense point cloud streaming with minor negative quality impacts compared to the baseline scenario when no adaptations are applied.
3. Virtual Transcendence Experience

Authors introduce the concept of Virtual Transcendence Experience (VTE) as a response to the interactions of several users sharing several immersive experiences through different media channels (cf. Figure 4). For that, authors review the current body of knowledge that has led to the development of a VTE system. This is followed by a discussion of current technical and design challenges that could support the implementation of this concept. This discussion has informed the VTE framework (VTEf), which integrates different layers of experiences, including the role of each user and the technical challenges involved. Authors conclude their paper with suggestions for two scenarios and recommendations for the implementation of a system that could support VTEs.

B. Volumetric

Volumetric media, popularly known as holograms, need to be delivered to users using both on-demand and live streaming, for new augmented reality (AR) and virtual reality (VR) experiences. As in video streaming, hologram streaming must support network adaptivity and fast startup, but must also moderate large bandwidths, multiple simultaneously streaming objects, and frequent user interaction, which requires low delay. In their paper, authors introduce the first system to our knowledge designed specifically for streaming volumetric media. The system reduces bandwidth by introducing 3D tiles (cf. Figure 5) and culling them or reducing their level of detail depending on their relation to the user's view frustum and distance to the user. Authors system reduces latency by introducing a window-based buffer, which in contrast to a queue-based buffer allows insertions near the head of the buffer rather than only at the tail of the buffer, to respond quickly to user interaction. To allocate bits between different tiles across multiple objects (cf. Figure 6), authors introduce a simple greedy yet provably optimal algorithm for rate-utility optimization. Authors introduce utility measures based not only on the underlying quality of the representation, but on the level of detail relative to the user's viewpoint and device resolution. Simulation results show that the proposed algorithm provides superior quality compared to existing video-streaming approaches adapted to hologram streaming, in terms of utility and user experience over variable, throughput-constrained networks.

C. Virtual Reality

Authors include methods and systems for streaming high-performance virtual reality video using adaptive rate allocation. In particular, an adaptive rate allocation system partitions a panorama video into segments or tiles and assigns priorities to each tile or segment based on input (e.g., a viewport of field-of-view) from a user client device. Further, the adaptive rate allocation system streams each tile or segment to the user client device according to the adaptive rate allocation, which maximizes bandwidth efficiency and video quality. In this manner, the adaptive rate allocation system delivers higher quality content to regions in the panorama video where a user is currently looking/most likely to look. In the VR (Virtual Reality) content, the immersion is important sensibility in order to be alive in the VR space. Also, the vision should be improved by following the sight of view and high quality of videos. For this reason, the VR content streaming uses a high-speed network for sending
continuously high frame rate and high frame size videos. There are many studies about video streaming technologies and QoS (Quality of Service) control mechanisms. However, they can’t be used for VR streaming technology in case of limited computer network resources. In their paper, authors introduce a QoS Management for Virtual Reality Contents by controlling QoS parameters according to user’s requests to keep the immersive experience quality in case of limited computer network resources.

1. 360° VR
Distributing high-quality 360° VR content to clients with modern networks using existing delivery techniques is economically challenging at scale. Despite the abundance of research in multihomed delivery and client-side adaptive bitrate (ABR) algorithms, current techniques continue to suffer from similar fundamental laws; namely the reliance on inflexible networks, inaccurate network models, and rigid control rules leading to suboptimal performance. Authors in their paper propose a unique approach to address many of these limitations. First, through the implementation of a modified version of an MPEG initiative called Server and Network Assisted DASH (SAND) for use in multipath networks enabling opportunistic messaging between the client, server, and network. Secondly, through the optimization of ABR algorithms that were generated using reinforcement learning and finally, by active management of multipath enabled transport protocols. Testbed experimental results revealed that our technique outperformed many state-of-the-art ABR algorithms in multipath networks. The average multipath differential delay decreased by 30% or more and outperformed standard Multipath TCP (MPTCP) and Quick UDP Internet Connections (QUIC) in key quality of experience (QoE) metrics by up to 20%.

D. Zoomable
Tiled streaming has been proposed for delivering ultra-high-resolution videos such as zoomable online lectures or panoramas. In tiled streaming, the source video is first partitioned into grid of small rectangular tile groups. Each tile group is independently encoded and compressed. When a user asks for a certain viewpoint at a time, the server only streams the viewed tiles to save up bandwidth. However, not much work has been done on finding the best tiling method for streaming panoramic video. Authors in their paper propose an effective tiling algorithm for tiled streaming by using both video content and user access preference distribution. Experimental results show that the proposed tiling method can save up to 32.4% and 69.8% of average streamed bitrate compared to conventional uniform tiling scheme and simply streaming the entire panorama respectively on equi-rectangular panoramic video.

E. Web 3D
Modern Web 3D technologies allow us to display complex interactive 3D content, including models, textures, sounds and animations, using any HTML-enabled web browser. Thus, due to the device-independent nature of HTML5, the same content might have to be displayed on a wide range of different devices and environments. This means that the display of Web 3D content is faced with the same Quality of Experience (QoE) issues as other multimedia types, concerning bandwidth, computational capabilities of the end device, and content quality (cf. Figure xx). In their paper, authors present a framework for adaptive streaming of interactive Web 3D scenes to web clients using the MPEG-DASH standard. Authors offer an analysis of how the standard’s Media Presentation Description schema can be used to describe adaptive Web 3D scenes for streaming and explore the types of metrics that can be used to maximize the user’s QoE. Then, authors present a prototype client they have developed, and demonstrate how the 3D streaming process can take place over such a client. Finally, authors discuss how the client framework can be used to design adaptive streaming policies that correspond to real-world scenarios.

---

**Algorithm: Proposed Tile-Growing Method**

1. Input: probability map $P$ and image frame $I$
2. Initialize tile map $T$ with $m \times n$ tiles
3. Do
4. Find all tiles including grown tiles in tile map $T$
5. For every tile $t_i$ do
6. Find all the nearest growing neighbors of tile $t_i$
7. Calculating $NGS$ of each neighbor by Eq. (5)
8. Find the largest $NGS$ as the best growing case
9. Judge if the largest $NGS$ is positive
10. End for
11. Compare $NGS$ of all the tiles in tile map $T$ and select the largest $NGS$ globally to update the tile map $T$
12. Until: tile map $T$ remains unchanged
13. Output: resulting tile map $T$

Fig. 7. DASH OMAF Architecture [6].

Fig. 8. Multipath SAND Testbed [6].

Fig. 9. The pseudocode of proposed tiling algorithm [26].
Video streaming is a foremost and growing contributor in the ever-increasing Internet traffic. Since last two decades, due to the enhancement in cameras and image processing technology, we have seen a shift towards multi-view plus depth (MVD) technology from traditional 2D and 3D video technology. This growth comes with deep changes in the Internet bandwidth, video coding and network technologies, which smoothed the mode for delivery of MVD content to end-users over the Internet. Since, MVD contains large amounts of data than single view video, it requires more bandwidth. It is a challenging task for network service provider to deliver such views with the best user’s Quality of Experience (QoE) in dynamic network condition. Also, Internet is known to be prone to packet loss, bandwidth variation, delay and network congestion, which may prevent video packets from being delivered on time. Besides that, different capabilities of end user’s devices in terms of computing power, display, and access link capacity are other challenges. As consequences, the viewing experiences of 3D videos may well degrade, if the quality-aware adaptation techniques are not deployed. In their article, authors work concentrates to present a comprehensive analysis of a dynamic network environment for streaming of 3D MVD over Internet (HTTP) [31]. Authors analyzed the effect of different adaptation of decision strategies and formulated a new quality-aware adaptation technique. The proposed technique is promoting from layer-based video coding in terms of transmitted views scalability. The results of MVD streaming experiment, using the proposed approach have shown that the video quality of perceptual 3D improves significantly, as an effect of proposed quality aware adaptation even in adverse network conditions.

1. Cloud 3D
Applying cloud technology to 3D interactive multimedia applications is a promising way to provide flexible and cost efficient online high bandwidth immersive services to a large population of end users [4]. One main reason cloud systems are popular among users is the fact that it relaxes the hardware requirements for high-end interactive visual applications. As most of the computational tasks are done on cloud servers, users no longer need to upgrade their hardware frequently to keep up with the ever-increasing high-end computing requirements of the latest applications. Moreover, cloud systems make it easier for a user to enjoy applications on different platforms, including mobile devices that are usually not powerful enough to run high-end, memory intensive services. In short, applying cloud technology to high end immersive applications has advantages in cost efficiency and flexibility both for the end users and the service providers. In previous work, authors proposed a flexible framework that addresses the bandwidth utilization and latency issues by using a combination of collaborative rendering, progressive meshes, and 3D image warping techniques. In their paper, authors analyze the performance of our framework for unreliable network environments where packets can be lost during network transmission. The experimental results show that compared to pure video streaming approaches, the proposed system is less sensitive to network conditions as part of the content is generated locally on the client.

F. 360° video
With 360° video, only a limited fraction of the full view is displayed at each point in time. This has prompted the design of streaming delivery techniques that allow alternative playback qualities to be delivered for each candidate viewing direction. However, while prefetching based on the user’s expected viewing direction is best done close to playback deadlines, large buffers are needed to protect against shortfalls in future available bandwidth. This results in conflicting goals and an important prefetch aggressiveness tradeoff problem regarding how far ahead in time from the current play point prefetching should be done. The authors of this paper present the first characterization of this tradeoff. The main contributions include an empirical characterization of head movement behavior based on data from viewing sessions of four different categories of 360° video, an optimization-based comparison of the prefetch aggressiveness tradeoffs seen for these video categories, and a data-driven discussion of further optimizations, which include a novel system design that allows both tradeoff objectives to be targeted simultaneously. By qualitatively and quantitatively analyzing the above tradeoffs, authors provide insights into how to best design tomorrow’s delivery systems for 360° videos, allowing content providers to reduce bandwidth costs and improve users’ playback experiences. The use of 360-degree videos to engage audiences in diverse contexts is increasing. While 360-degree videos have the potential to create new value in enhancing audiences’ viewing experiences, they often decrease audience engagement by causing motion sickness in an immersive environment. Despite increasing scholarly and practical attention to the effect of 360-videos on audience engagement, the question of how to enhance it through immersive 360-degree videos remains unanswered. Therefore, this study empirically examined the effects of different display types and viewport dynamics on audience engagement using data collected from 60 subjects during a laboratory experiment. The results show that an audience’s viewing experience in an immersive environment is influenced by the joint effects of display types and viewport dynamics. By explaining the mechanisms by which audiences are engaged with 360-degree videos, this
study contributes to resolving previous inconsistent findings regarding the effect of immersive technology on audience engagement.

Fig. 11. Example of personalized layers based on viewing direction and downloaded tiles [2].

G. Imaging
Today, an ever-increasing number of diseases requiring a surgical solution are treated by means of minimally invasive procedures [1]. When possible, these procedures are suggested instead of traditional approaches because they decrease both the surgical risk and postsurgery hospitalization. The growing technological advances in terms of imaging techniques (i.e., visualization of organs and tissues) and handling of the minimally invasive instruments (e.g., endoscopes) play a key role in improving the qualitative aspects of the interventions [2].

Fig. 12. The logical architecture of the proposed IVRE system [3]

IV. CONCLUSION
In recent years DASH has become the de facto standard for video streaming. However, applications continue to demand more bandwidth in an effort to satisfy users. This is based on applications such as 360° and immersive VR/AR video. This paper explores current state of the art and introduces the reader to some of these DASH-aware bandwidth intensive applications. We break applications down into main categories: 360°, Volumetric, Virtual Reality (VR), Imaging, Immersive, Zoomable and Web 3D. Further, Volumetric is broken down into Augmented Reality (AR) 3D and VR. VR is broken down into VR 3D, VR 2D and 360° VR. Immersive is broken down into Immersive 3D and Virtual Transcendence Experience (VTE). Finally, Cloud 3D is a component of Web 3D. An outline of these applications is given. Authors anticipate that most future work on DASH-based applications will be in the areas illustrated in this paper.

REFERENCES


Author Details

Koffka Khan received the M.Sc., and M.Phil. degrees from the University of the West Indies. He is currently a PhD student and has up-to-date, published numerous papers in journals & proceedings of international repute. His research areas are computational intelligence, routing protocols, wireless communications, information security and adaptive streaming controllers.

Wayne Goodridge is a Lecturer in the Department of Computing and Information Technology, The University of the West Indies, St. Augustine. He did is PhD at Dalhousie University and his research interest includes computer communications and security.