A Literature Survey on Adaptive Streaming in Heterogeneous Networks

Namitha Murugesh
Dept of Computer Science and Engineering
B M S College of Engineering
Bangalore, India
namithasendil79@gmail.com

ABSTRACT

A growing number of mobile users require uninterrupted audio/video streaming while they are on the move. They move through various heterogeneous networks. They expect seamless mobility as they move with their devices equipped with multiple interfaces. To support the continuous provisioning of multimedia flows i.e., service continuity during handoffs when a wireless device disconnects from one network and re-connects to a new one is a challenging task. In particular vertical handoff situations that occur when mobile devices dynamically change not only their access points but also the wireless access technology/infrastructure they are using, e.g., from WiFi to 3G possibly requiring dynamic content adaptation. This survey concentrates on various dynamic content adapting techniques, protocols that aid in streaming, handoff procedures that affect continuity in streaming and proxy nodes that are used to maximize the QoS and experience of user.

Keywords : Content adaptation; ubiquitous; heterogeneous networks; streaming; vertical handoff

1. INTRODUCTION

The 4G technology promises faster data rates opening up avenues for application that needs high data rates.like video streaming The next generation networks promise to bring together heterogeneous networks like Wi-Fi, Wi-Max and cellular networks to make it one ecosystem. All of these networks will be treated as one family of packed switched networks. The demands for streaming video, Internet radio and other internet access has increased in the recent. The user has the luxury of roaming through the heterogeneous networks while the streaming is going on. The handoffs take place between the different networks. When the network characteristics are different the handoff is called a vertical handoff. To facilitate streaming while the user roams through different networks, the content has to be adapted to maintain the streaming presentable to the user. The adaptation process must be capable of maintaining the streaming content's quality to ensure that the handoff seems as seamless as possible to the user. The adaptation of the content is done according to the transmission characteristics of the end-to-end communication path and to the capabilities of the displaying device.

Adaptation is a process which repackages the content being streamed according to the present eco-system characteristics. Here the eco-system consists of the end-user device, network characteristics, content requested for streaming and the intermediate nodes like proxy. Different combinations of these entities would affect the QoE (Quality of Experience) of the user. The end user's device's screen size, computational capacity, battery power, available bandwidth, loss due to wireless network play an important role in deciding the final version of the content viewed by the user. In an ongoing streaming application when a user roams from one network to another network the network parameters change and the session has to be transferred seamlessly. When the user wanders through heterogeneous networks vertical handoff takes place. The transfer of session must be done seamlessly such that the user does not experience an interruption with the service. The efficiency of the adaptation system is many ways related to the handoff procedures. The combinations of both activities play a vital role in the QoE of the user. Researchers in Ref. [1] explain how QoE could be improved by both network QoS and application QoS management.

The survey focuses on the different technologies, mechanisms and protocols that help to adapt the content being streamed from the source to the client in a seamless way as the client roams around in a heterogeneous network. Section 2 describes the codecs and video formats that aid in content adaptation. Section 3 explores the role of proxy in content adaptation. Section 4 describes the different transport protocols that are used for streaming. Section 5 explores the different handoff procedures.

2. CONTENT FORMATS AND THEIR IMPACT ON ADAPTATION

Adaptation is carried out at the different layers of the protocol stack. This research is concentrating on adaptation techniques at the application layer . Even though the adaptation process happens at the application layer the feedback from the transport layer plays a crucial role.

Transcoding is a process where the content is re-coded into a new format. The transcoded video streams can have a lower spatial resolution, a lower temporal resolution, a lower quality, or even a different compression standard [2]. The new format is decided based on the target devices capability and constraints that are present in the network connection. In transcoding the content is decompressed completely or to an intermediate form and again recoded into a form that is decodable by the client's device. A single source sequence is kept in the video storage and different versions are created on-the-fly upon request using transcoding methods. An intermediate solution providing transcoding at a low complexity by the aid of control streams is proposed in [3].

The Scalable Video Coding (SVC) as in H.264 [4],[5] was developed in response to the growing need for higher compression of moving pictures for various applications such as videoconferencing, digital storage media, television broadcasting, Internet streaming, and communication. It is also designed to enable the use of the coded video representation in a flexible manner for a wide variety of network environments. Here in SVC the content is encoded once and can be decoded in several layers to suit the requirements of the target device and network conditions. It is a coding standard in which the video is coded with a base layer video stream meant for connections with basic terminal capabilities or low bandwidth network conditions.

The residual information between the base layer and the original content is then encoded to form one or more enhancement layers. Additional enhancement layers can be integrated with the base layer for scaling up the quality of stream. Thus giving the user the flexibility to choose the quality of stream that can be received. SVC extension of the H.264/AVC standard has achieved significant improvements in coding efficiency with an increased degree of supported scalability relative to the scalable profiles of prior video coding standards. MPEG-4/AVC outperformed MPEG-2 in terms of throughput, packet delays, packet loss and jitter. Performance of mobile video streaming in different scenarios depending on the user's movement speeds and video coding standards were presented in [6]. The authors in [7] propose a dynamic adaptation scheme of SVC to optimally adapt video stream over heterogeneous networks using the MPEG-21 Digital Item Adaptation (DIA) tool. MPEG-21 DIA framework provides systematic solutions in choosing the optimal adaptation operation to given conditions and supports interoperable video adaptation. Their experiment results showed that the proposed adaptation scheme provides QoS-enabled delivery and consumption of SVC with time-varying constraints of network, terminal, and user preference, in a robust and efficient way. In [8] Razib Iqbal and Shervin S show that the adaptation operations can be quicker, when adaptation systems are designed to adapt contents according to the encoding structure but in an intermediary node following a codec-independent technique. The adaptation is performed on-demand based on its generic Bitstream Syntax Description (gBSD). Their approach was to avoid conventional cascaded or multiple pre-coded bitstream adaptations.

To sum it all up, the major downside of transcoding is the additional complexity needed to re-encode the video sequence in its new form. Scalable coding is less efficient compared to single layer coding when one fidelity version of the video stream should be transmitted. The layering is an unnecessary overhead. On the other hand layered coding gives flexibility in choosing the quality of content dynamically which plays an important role in seamless continuous streaming.

Bit-rate Adaptation is one of the popular techniques to adapt the content. Authors in [23] propose a rate adaptation algorithm which detects bandwidth changes based on segment fetch time. They use a smoothed throughput measure rather than an instantaneous measure. The advantage to this method is that it does not require the information from the transport layer since the calculation is done at the application layer. Though it saves time in the cross layer message exchange the adaptation process might be more accurate if the information from the transport layer is used. As in this research we would like to know if the bandwidth change is due to an anticipated handoff. If a handoff occurs the characteristics of the new network have to be taken into account. Another technique is to use the client's buffer status as a metric to adapt the rate at which the content can be received. The content quality is directly related to the data size that has to be received and decoded by the client. Whether the buffer faces underflow because of a slow connection or faces an overflow since the arrival rate exceeds the buffer capacity can be addressed by a feedback mechanism which controls the data rate. Chenghao Liu et al [24] propose a sub-stream level of a client buffer feedback , scalable streaming adaptation based on Multiple Virtual Client Buffer Feedback (MVCBF) and Target Virtual Buffer Protection Time (TVBPT). They use RTSP protocol in their experiment which is a contrast to the popularity of HTTP in adaptive streaming. They show that their method outperforms the scalable rate adaptation method based on client buffer feedback as specified in 3GPP PSS. To improve the quality of experience for the users, 3GPP

Packet-Switched Streaming Service (PSS) [25] supports the adaptation of continuous media. When rate adaptation is performed at the server, the feedback from the client to the server is used at the server to adjust the transmission bit rate and the quality in an end-to-end scenario. In order to signal the client feedback information from the client to the server, 3GPP introduced a new Application-defined Real Time Control Protocol (RTCP) called Next Application Data Unit (NADU) that is delivered as a report block together with other RTCP feedback information. In the literature, a comprehensive discussion of adaptive streaming within the 3GPP packet-switched streaming service has been presented in [26]. Kampmann et al. [27] proposed a 3GPP PSS compliant stream switching solution. In the proposed solution, stream switching is carried out based on certain thresholds of the total media time in the client buffer and using NADU feedback at the server. Schierl et al. [28] presented a 3GPP PSS compliant adaptive video streaming strategy for H.264/AVC encoded stream. Here, a transmission rate estimation algorithm was proposed based on NADU and RTCP Receiver Report (RR) from the client to the server. Based on the estimated transmission rate, adapting the video bit rate is realized by the combination of Bit-Stream Switching and Temporal Scalability. However, all those adaptive video streaming solutions are carried out based on a single layer client buffer feedback.

3. PROXY

An intermediate node between the server and the client identified as appropriate for performing the content adaptation, store and forward is called the proxy. The proxy or gateway, receives instructions from the receiver prior to the stream's initiation regarding the parameters of the adaptation process. For the mobile devices featured with lower bandwidth network connectivity, transcoding can be used to reduce the object size by lowering the quality of a multimedia object. In view of the monolithic transcoders which only provide transcoding services and have limited performances due to the unknown data types and protocols in the prior research the authors of [20] propose the architecture of versatile transcoding proxy (VTP). Based on the concept of the agent system, the VTP architecture can accept and execute the transcoding preference script provided by the client or the server to transform the corresponding data or protocol according to the user's specification. Media cloud services offer a unique opportunity for alleviating many of the technical challenges faced by mobile media streaming, especially for applications with stringent latency requirements. A novel cloud-assisted architecture is proposed in [21] for supporting low-latency mobile media streaming applications such as online gaming and video conferencing. A media proxy at the cloud is envisioned to calculate the optimal media adaptation decisions on behalf of the mobile sender, based on past observations of packet delivery delays of each stream. The proxy-based intelligent frame skipping problem is formulated within the Markov Decision Process (MDP) framework, which captures both the time-varying nature of video contents as well as bursty fluctuations in wireless channel conditions. The optimal frame skipping policy is calculated using the stochastic dynamic programing (SDP) approach, and is shown to consistently outperform greedy heuristic schemes. In general, the sizes of multimedia files are much larger than those of regular webpages. It is unlikely that a steaming proxy server can constantly store entire contents of multimedia files in its memory. As a result, the streaming proxy server needs to split individual multimedia files into segments and only store popular segments in its memory. Researchers had proposed various ways to do the segmentation. However, the sizes of individual segments are often fixed once the segmentation is done. Tsozen Yeh and Zongwei Yang [22] argue that the sizes of individual segments should vary according to their popularity. A popular segment can have a longer length so the overall performance can be increased accordingly. A novel design, Dynamic Segment Size (DSS), which dynamically adjusts the length of segments by their popularity, is proposed.

The design is applied to a sophisticated algorithm, Adaptive and Lazy Segmentation (ALS), which performs the work of splitting multimedia files into segments and handling memory replacement in a streaming proxy server. The advantages that come with the adoption of the proxy solution is that it can be located at the most critical position in the end-to-end path. The complexity of the proxy architecture is significantly higher and requires gateways with powerful CPUs and a lot of memory. The degree of the receiver□s participation in the adaptation process can dictate the applicability and the effectiveness of the proxy adaptation scheme.

4. TRANSPORT PROTOCOLS

The content requested for can reach the client by different transport protocols. The popular protocols are RTSP, RTMP and HTTP. RTSP is specifically designed to be used for delivering streaming media. Trick-play modes such as fast-forward or rewind, using VCR-like controls are supported with RTSP unlike HTTP which works best when segments are sent in sequential order. Viewing can also begin the moment the first bits reach the RTSP player; meaning that a 2- or 10-second segment delay doesn't affect RTSP delivery. Adobe uses a proprietary messaging protocol called RTMP (Real-Time Messaging Protocol) for its delivery from Flash Media Server (FMS) to user's Flash Player in-browser playback. It is a variant of RTSP. In multicasting scenarios RTSP can support multicasting by delivering a single feed to many users, without having to provide a separate stream for each of them. HTTP is a true one-to-one delivery system. RTMP like RTSP is defined as a protocol that saves the state of session. From the first time a client player connects until the time it disconnects, the streaming server keeps track of the client’s actions or “states” for commands such as play or pause. When a session between the client and the server is established, the server begins sending video and audio content as a steady stream. This behaviour continues and repeats until the server or player client closes the session. Recent advancements also accommodate for potential brief interruptions in the server-client connection, allowing for a small amount of content to be played back from a local buffer. Encryption is another hallmark of RTMP, as RTMP encrypted (RTMPE) protects packets on an individual basis (more on this later). Some HTTP-based solutions are beginning to address integrated digital rights management (DRM), but the majority of HTTP delivery cannot support encryption at a packet level. The tunnelling feature in RTMP called RTMPT allows RTMP to be encapsulated within HTTP requests. This allows RTMP to traverse firewalls by appearing to be HTTP traffic on Port 80. HTTP streaming has gained popularity in recent years for the following reasons. Larger segments of multimedia can now be delivered efficiently using HTTP. Support for HTTP in the present Internet infrastructure is favorable. CDNs have evolved to serve multimedia services. HTTP is firewall friendly because almost all firewalls are configured to support its outgoing connections. HTTP streaming is light on the server since the client manages the streaming without having to maintain a session state on the server. The popular streaming platforms that use HTTP streaming as their underlying delivery method are Apple's HTTP Live Streaming[9], Microsoft's Smooth Streaming[10] and Adobe's HTTP Dynamic Streaming[11]. However, each implementation uses different manifest and segment formats and therefore to receive the content from each server, a device must support its corresponding proprietary client protocol. A standard for HTTP streaming of multimedia content would allow a standard-based client to stream content from any standard-based server, thereby enabling interoperability between servers and clients of different vendors. MPEG-Dynamic Adaptive Streaming (DASH) [12] is being developed to facilitate the idea of a common ecosystem of content and services that will be able to provision a broad range of devices such as PCs, TVs, laptops, set-top boxes, game consoles, tablets, and mobiles phones. The multimedia content is captured and stored on an HTTP server and is delivered using HTTP. The content exists on the server in two parts: Media Presentation Description (MPD), which describes a manifest of the available content, its various alternatives, their URL addresses, and other characteristics; and segments, which contain the actual multimedia bitstreams in the form of chunks, in single or multiple files. To play the content, the DASH client first obtains the MPD. The MPD can be delivered using HTTP, email, thumb drive, broadcast, or other transports. By parsing the MPD, the DASH client learns about the program timing, media-content availability, media types, resolutions, minimum and maximum bandwidths, and the existence of various encoded alternatives of multimedia components, accessibility features and required DRM, media-component locations on the network, and other content characteristics. Using this information, the DASH client selects the appropriate encoded alternative and starts streaming the content by fetching the segments using HTTP GET requests. After appropriate buffering to allow for network throughput variations, the client continues fetching the subsequent segments and also monitors the network bandwidth fluctuations. Depending on its measurements, the client decides how to adapt to the available bandwidth by fetching segments of different alternatives (with lower or higher bitrates) to maintain an adequate buffer. The MPEG-DASH specification only defines the MPD and the segment formats. The delivery of the MPD and the media-encoding formats containing the segments, as well as the client behaviour for fetching, adaptation heuristics, and playing content, are outside of MPEG-DASH□s scope. Streaming paths in heterogeneous networks include IEEE 802.11 wireless as well as 3G/4G mobile networks. TCP's shortcomings on wireless paths degrade the performance and QoS for high definition media streaming while the actual bandwidth provisioning on those networks is no longer the limiting factor. Due to physical packet loss and large propagation delay the transport protocol suffers from significant underutilization of the available bandwidth. For dynamic HTTP streaming this underutilization translates directly into unnecessary quality reduction. In order to improve the quality of dynamic streaming on wireless networks, Manuel Gorius, Yongtao Shuai and Thorsten Herfet [13] implement a novel transport protocol - Predictably Reliable Realtime Transport (PRRT), a protocol layer that efficiently supports the reliability required by multimedia services under their specific time constraint. The dynamic adaptive streaming also will be benefited with this approach.

5. HANDOFF-PROCEDURES

When the mobile user wanders out of the coverage area of the present network and into coverage area of a different wireless network a handoff has to take place between the networks. The handoff can be a homogeneous also known as horizontal handoff if the two networks have similar characteristics e.g. between two Wi-Fi spots. If the characteristics of the networks are different it is called a heterogeneous or vertical handover e.g. form Wi-Fi to 4G cellular network. The techniques used for managing handover can be classified depending on the layer of the network stack at which the handover is done. The possible classes are handover at network, transport or application layer. Mobility at the network layer is provided by Mobile IP [14]. This class of handoff is efficient regarding the period the stream is interrupted on down side the multimedia session cannot be adapted to the parameters of the new access network. Also MIP presents poor scalability and high packet loss. The transport layer handoff continues the transport connection during the network switch while changing the associated IP address. Transport layer mobility management is achieved through the Mobile Stream Control Transmission Protocol (mSCTP) [15]. mSCTP resolved the problem of packet loss encountered at network layer mobility by pausing the transmission during mobility induced disconnections. mSCTP offers an efficient handover management by using optimization of the path-transition and failover mechanism. The disadvantage of mSCTP is that it does not address context aware quality of service. Several mechanisms for mobility support at application layer have also been defined. If the handover is triggered before losing connectivity, the application layer mobility performs better than the other two solutions. This approach offers another important advantage, because the multimedia session parameters can be adapted to the available resources from the connected network. The specifications for application layer mobility were defined by 3GPP working group. They developed a multimedia session continuity (MMSC) [14] framework for session transfer between packet switched networks (PS) in conjunction with the session transfer between packet switched networks and circuit switched network (CS), standardized already as Voice Call Continuity (VCC). Elena Apostol and Valentin Cristea [15] propose architecture that handles multimedia streaming adaptation, session management and takes into consideration user profiles and a set of quality of service criteria. To ensure session continuity when a user moves from one network to another this architecture supports handover at the application layer which offers a better quality control than the majority of mobility solutions. Session continuity during handover by integrating mSCTP into the 3GPP IMS [18] architecture is presented in [19]. IMS is the service control layer to add QoS control and context-aware functions.

6. CONCLUSION

Though there are numerous research efforts in adapting streaming media according to the network characteristics and end-user equipment. Seamless transfer of the on-going session as the client wanders through heterogeneous networks is still a challenging process. Here we have surveyed the different components of the framework required to achieve a seamless content adaptation during a streaming service.

7. REFERENCES

[1] Ricky K. P. Mok, Edmond W. W. Chan, and Rocky K. C. Chang, “Measuring the Quality of Experience of HTTP Video Streaming,” IEEE International Symposium on Integrated Network Management (IM), 2011 IFIP, 2011,pp. 485-592.

[2] Amon. P, Haoyu Li, Hutter. A, Renzi. D ans Battista. S, “Scalable Video Coding and Transcoding, “IEEE International Conference on Automation, Quality and Testing, Robotics, pp. 336-341, May 2008.

[3] Glenn Van Wallendael, Jan De Cock, and Rik Van de Walle,” Fast Transcoding For Video Delivery By Means Of A Control Stream,” Conference on Image Processing (ICIP), 2012 19th IEEE International, pp. 733-736, 2012.

[4] Advanced video coding for generic audiovisual services, ITU-T-REC- H.264, ITU-T,Jan-2012.

[5] Heiko Schwarz, Detlev Marpe and Thomas Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 17, no. 9, September -2007.

[6] Saleh Abdallah-Saleh, Qi Wang and Christos Grecos,” Evaluation of Mobile Video Streaming in Heterogeneous Wireless Networks,” Telecommunications Forum (TELFOR), 2011 19th, pp. 262-265 ,November-2011.

[7] Haechul Choi, Jung Won Kang and Jae-Gon Kim, “Dynamic and interoperable adaptation of SVC for QoS-enabled streaming,” IEEE Trans. on Consumer Electronics, vol.53, pp. 384-389, May 2007.

[8] Iqbal, R. and Shirmohammadi, S, “MPEG-21 based temporal video adaptation for heterogeneous devices and mobile environments,” IEEE International Conf. on Multimedia and Expo, 2009. ICME 2009, pp. 1845-1846, 2009.

[9] R. Pantos and E.W. May, ‘‘HTTP Live Streaming,’’ IETF Internet draft, work in progress, Mar. 2011.

[10] Microsoft, IIS Smooth Streaming Transport Protocol, Sept. 2009; http://www.iis.net/community/files/media/smoothspecs/[MS-SMTH].pdf.

[11] http://www.adobe.com/products/httpdynamicstreaming

[12] Sodagar, I, “The MPEG-DASH Standard for Multimedia Streaming Over the Internet,” MultiMedia, IEEE, vol. 18, pp. 62 – 67, April-2011.

[13] Gorius. M, Yongtao Shuai and Herfet. T, “Dynamic media streaming over wireless and mobile IP networks,” IEEE International Conference on Consumer Electronics - Berlin (ICCE-Berlin), pp. 158-162, September 2011.

[14] Xinyi Wu and Gang Nie, “Design and Simulation of an Enhanced Handover Scheme in Heterogeneous Mobile IPv6 Networks,” Conference on Information Processing, pp. 448-451, July 2009.

[15] Łukasz Budzisz, Ramon Ferrus and Ferran Casadevall, “Design principles and performance evaluation of mSCTP-CMT for transport-layer based handover,” Vehicular Technology Conference,pp. 1-5, 2009.

[16] The Third Generation Partnership Project, Feasibility Study on Multimedia Session Continuity. VCC Release 8, June 2008,

[17] Elena Apostol and Valentin Cristea, “Multimedia Mobility Service across Heterogeneous Environments,” International Conference on Emerging Intelligent Data and Web Technologies, pp. 172-177, September 2011.

[18] TS 23.328, “IP Multimedia Subsystem,” 3GPP, Release 6.

[19] Nguyen Huu Thanh, Le Thi Hang, Ngo Quynh Thu, Vu Van Yem, Nguyen Xuan Dung, “Multimedia Session Continuity with Context-Aware Capability in IMS-based Network,” International Symposium. on Wireless Communication Systems, pp. 383-387, September-2009.

[20] Jung-Lee Hsiao, Hao-Ping Hung and Ming-Syan Chen, “Versatile Transcoding Proxy for Internet Content Adaptation,” IEEE Transactions on Multimedia, pp. 646- 658, 2008.

[21] Xiaoqing Zhu, Jiang Zhu, Rong Pan, Prabhu, M.S and Bonomi. F, “Cloud-assisted streaming for low-latency applications,” International Conference on Computing, Networking and Communications (ICNC), pp: 949- 953, 2012.

[22] Tsozen Yeh and Zongwei Yang, “Using dynamic segmentation adjustment to improve the performance of streaming proxy servers,”IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1 – 5, 2012.

[23] Liu, Chenghao, Bouazizi, Imed and Gabbouj, Moncef, "Rate Adaptation for Adaptive HTTP Streaming,"Proceedings of the second annual ACM conference on Multimedia systems,pp. 169-174, 2011.

[24] Liu, Chenghao, Bouazizi, Imed and Gabbouj, Moncef, "Advanced rate adaption for unicast streaming of scalable video," 2010 IEEE International Conference on Communications (ICC), pp. 1-5, 2010.

[25] 3GPP TS 26.234, “Transparent End-To-End Packet-Switched Streaming Service (PSS): Protocols and Codecs (Release 6) P. Fröjdh, U. Horn, and M. Kampmann, “Adaptive Streaming within the 3GPP Packet-Switched Streaming Service,” IEEE Network, vol 20, no 2, pp. 34-40, March-April 2006.

[26] M. Kampmann, and C. Plum, “Stream switching for 3GPP PSS compliant adaptive wireless video streaming,” IEEE Consumer Communications and Networking Conference, 2006 (CCNC 2006), Jan. 2006.

[27] T. Schierl, M. Kampmann, and T. Wiegand, “3GPP Compliant Adaptive Wireless Video Streaming Using H.264/AVC,” in IEEE Int. Conf. Image Proc., Genova, Italy, Sept. 2005.