Technological Developments in Video-Streaming Applications

Pragna G S [1], Roopa J [2]

[1]*Electronics and Communication Dept., R.V. College of Engineering, Bangalore, India*
[2]*Assistant Professor, Electronics and Communication Dept., R.V. College of Engineering, Bangalore, India*

Abstract - *Streaming video services over the internet has been rapidly increasing over the past few years with the advancement in networking technology. Internet-streaming provides a lot of advantages over the traditional cable television, which is the reason why it has gained popularity, and has even become a necessity. There is a requirement for a well-developed architecture supporting and enabling the process of video-streaming over the internet. This paper provides a review of some of the essential technological requirements for internet-based video-streaming and how they help to make the experience of the user better, such as Content Delivery Networks, Adaptive Bitrate Streaming, video compression techniques, Quality of Experience, and development of user-interface.*

Key Words: **video-streaming, content delivery networks, adaptive bitrate streaming, video compression, quality of experience, OTT**

1. INTRODUCTION

Over-the-top (OTT) media services and live-streaming over the internet, have seen an increasing audience over the past few years with the technological advancement in network, connectivity, platform-development, etc. There is a growth in the general ease of use, since people find it convenient to stream the content on almost any device with an internet connection, and are provided access to a wide range of content from around the world which they are free to watch whenever they want. This is largely different from traditional media service providers such as satellite television, cable and broadcast platforms (Connected TV or CTV). The difference is that, in internet-streaming, the transfer of content happens over the internet, and the data is divided into packets or chunks, whereas traditional television require setting up of cables, set-top box, antenna, converter block, etc. which means that they cannot be mobile. In cable TV, the service is limited in terms of its quality and resolution, but internet-streaming has the feature of adaptive streaming, which means that it is accessible by a wide range of devices, network connectivity, and resolution. Cable TV broadcasts are generally restricted to certain regions, but internet-streaming is necessary for a global delivery of streaming services. One important advantage of cable TV however, has been that the latency during live-streaming is extremely less, whereas in internet-streaming it is generally more comparatively (sometimes up to 45 seconds). However, a few new protocols such as low-latency HLS (HTTP Live Streaming) and low-latency CMAF (Common Media Application Format) for DASH (Dynamic Adaptive Streaming over HTTP) have reduced the end-to-end delivery time to about less than 3 seconds.

Since there is an increasing demand for video-streaming services, it is extremely necessary that the architecture that supports it must be very strong and rapidly developing and advancing with new and upcoming technology. To make such streaming possible at a large scale, it is essential to have certain features implemented, and this paper reviews some of those essential features that support today's video streaming applications. These features include the implementation of Content Delivery Networks, Adaptive Bitrate Streaming, advanced video compression techniques, giving importance to Quality of Experience of the user, and development of User Interface.

2. TECHNOLOGICAL FEATURES

Some of the technological features that help provide better streaming services are enlisted below and explained.

2.1 Content Delivery Networks

Speed of content delivery and response time are of utmost importance to video streaming services. Due to increased user demand, there is also a lot of load on the servers. Content Delivery Networks (CDN) act as edge servers and provide an efficient management system to help in distribution and routing of digital content. They are located around the globe, so that the end users can access the CDN which is nearby to them, thereby reducing the network traffic on the original servers. A general representation of CDN is shown in Fig. 1.

As compared to a network in which the original server serves all the users, CDN acts as a mediating entity between the original server and the users.

The most important strategies involved in designing a CDN are content caching and the location of the server. Files are cached by the CDN near the user's location, which increases the response speed and reduces delay. CDN nodes work in cooperation with each other and it is essential to decide which node stores what kind of data, and this is usually determined by the demands of the user in that particular location [1]. Multiple intermediate nodes can be added capable of caching and multiplexing capabilities. CDN nodes are designed to support an edge server compared to network devices whose hardware is built for routing and switching [2].

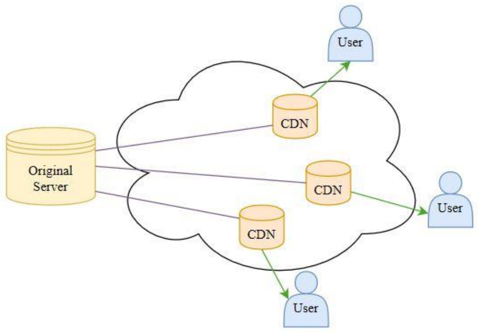

Fig -1: Representation of Content Delivery Network (CDN)

The advantages of caching are that it reduces bandwidth and reduces latency. Both the ideal and practical hit rates of caching in CDN are generally high compared to the normal web servers [3]. In order to improve caching performance, newer algorithms are being adopted to select the nodes which have to be cached based on the closeness, reach, degree, or betweenness centrality of the nodes. Cooperative edge caching makes use of a shared storage between base stations and they work in collaboration to choose the caching strategy and cache even the popular files that are not cached by other nodes. It is also necessary to employ mobile edge caching techniques, since most of the devices connected to the internet are mobile devices. This means that the content dynamics is rapidly changing and there is an increased chance of cache miss. In such cases, artificial neural network learning algorithms need to be employed to observe and learn the real-time requirements [1].

2.2 Adaptive Bitrate Streaming

The Adaptive bitrate (ABR) streaming is a method used to deliver the video to the user in the most efficient way possible and in the highest usable quality for each specific user. If there is no adaptive streaming, there will be a lot of buffering if the user has poor internet connection and there will be problems with video quality and resolution. Adaptive streaming allows the video provider to create different videos for different screen resolutions and bitrates, allowing the end user to automatically choose the appropriate video corresponding to their network bandwidth and screen resolution.

Fig. 2 shows a simple representation of how adaptive bitrate streaming is implemented. The video provider produces a video content with high bitrate and good resolution. This is passed through an encoder where the file is converted into multiple bitrates and stored as separate files in the server. The list of these available stream segments and their bitrates are stored in the manifest file, which the user-interface (UI) downloads first. The UI then starts downloading the video content depending on the lowest available network bandwidth and the default resolution. When the network connectivity improves, it can choose between the different encoded files and stream the one with a better bitrate in chunks. The chosen encoded file can be dynamically changed to suit the real-time requirement, it can be client-driven, or even sometimes server-driven or done at the network level.

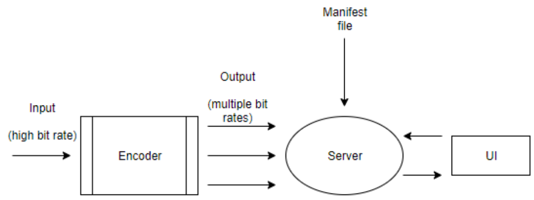

Fig -2: Adaptive Bitrate Streaming

The factors within the scope of the ABR algorithm which affect the performance of content selection include resource related factors (such as throughput, buffer and power), scheduling factors and the technique of adaptive selection (heuristic, AI, etc.) as well as external factors such as the nature and size of the content and the performance of the CDN [4]. It is important for the ABR algorithms is to prevent buffer overflow which maintaining the Quality of Experience (QoE) for the users [5].

The MPEG-DASH, developed under the Moving Pictures Experts Group (MPEG), is a standard developed to provide uninterrupted and smooth streaming services using principles of ABR algorithms. In DASH, video data is received in the form of chunks rather than a continuous stream. To describe information regarding the URL that needs to be fetched, list of bitrates and resolution, timing, etc., a media presentation description (MPD) file is used. Another widely used technique is HLS developed by Apple Inc.

2.3 Video Compression Techniques

Recording of video content is done at high resolutions, so that they can later be rendered at a low or high quality as per the demands of the client. But transmission of this high-quality content over the network is not feasible. Hence there is a necessity for efficient video compression techniques. These techniques have to reduce the size of the video by removing the redundant information such that the encoded file consists of as much as useful information as possible with the least amount of memory footprint. If compression is more, the quality will be less but data reduction will be more.

Motion estimation is the technique by which the redundancy between two successive frames is reduced by comparing the previous and current frames, and storing only the part in the frame that has changed over time. The technique of predicting and compensating for the relative change in motion between two frames is called as motion compensated prediction. Discrete Cosine Transform (DCT) and other forms of predictive and transform coding are used to reduce spatial redundancy. Color space inversion is used to reduce color space redundancy [6].

Lossy compression techniques are classified into two types, the intra-frame compression and inter-frame compression. In intra-frame compression, each frame is compressed as a separate image by making use of spatial redundancy and identifying the correlation of pixels within that frame. It usually requires less processing time. In inter-frame compression, successive frames are compared making use of temporal redundancy. More processing time is required since a set of frames need to be compared with respect to one another.

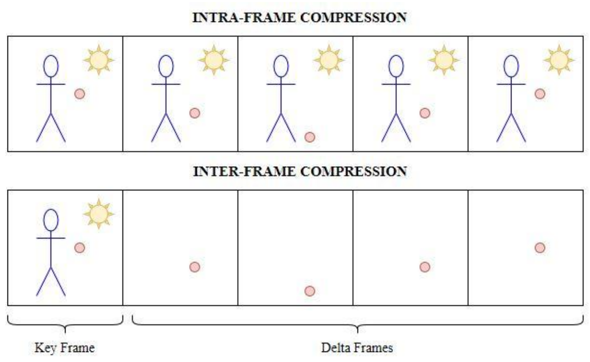

Fig -3: Representation of intra and inter-frame compression

A representation of the comparison between intra-frame and inter-frame compression techniques is shown in Fig.3. In intra-frame compression all the components of the frame are stored by reducing some spatial redundancies with respect to neighboring pixels. In inter-frame compression, the initial frame is taken as the "key frame" and all the components of it are stored. In the upcoming frames, only the ball is moving but the background involving the image of the person and the sun are staying static. Hence only the small portion of the frame containing the ball is stored in smaller "delta frames" and all the other parts of those frames are discarded. While decoding, the key frames are attached to each of the delta frames and rendered to the user. This ensures that a maximum portion of the screen maintains its resolution and quality.

Some of the well-known compression techniques from MPEG include MPEG-1, MPEG-2, MPEG-4, H.261, H.263, H.264, H.265, and H.266. Most of the video encoding makes use of H.264 Advanced Video Coding (AVC), which is suitable for streaming on most devices. The H.265 High Efficiency Video Coding (HEVC) supports 8K resolution and is suitable for streaming high quality content, but is not widely used due to royalties. In order to avoid the complication of royalties, codecs such as AV1 and VP9 have been developed, which are open-source and royalty-free alternatives. The upcoming H.266 Versatile Video Coding (VVC) aims to provide better compression that its predecessor, but comes with royalty issues. Video streamers often make use of multiple codecs for encoding and decoding so that it can be delivered to a large number of users.

2.4 Quality of Experience

There is a rapidly growing demand for video content streamed over the internet services. There is a requirement to have a standard of measuring the experience of the user while using the service and ensure whether the expectations are met or not.

Quality of Service (QoS) is used to measure network parameters such as latency, packet loss rates, throughput, etc, but Quality of Experience (QoE) measures the actual experience of the end-user, and whether the delivery was satisfying enough. QoS is network-centric, whereas QoE is user-centric. It is a measure of whether the user is happy or frustrated while viewing the content, and if the overall quality was enjoyable or not.

The metrics to be considered while measuring QoE can be classified as objective metrics and subjective metrics which depend on video codecs and network characteristics [7]. Objective metrics include Bitrate (adaptive bitrate provides a good experience, but the bitrate should not fluctuate often), Peak Signal-to-Noise Ratio (PSNR) and Video Quality Metrics (VQM) which is carried out between the original uncompressed video and the reconstructed video. Subjective metrics include playback smoothness (such as delay, jitters, discontinuous playback, etc. as perceived by users) and employing an observer test-bench to review and evaluate the performance and assess the quality.

The factors that influence QoE are termed as Influence Factors (IF) and are of types context IF, system IF, and human IF [7]. Context refers to the conditions and situations at which the end-user is accessing content, such as physical context (location of user, their movement, etc.), temporal context (duration for which content is played), or economic context (type, cost of device, subscription service). System IF refers to the technical properties of the supporting system or architecture, which may be content related (compression technique used), network related (bandwidth, latency, throughput, congestion control, etc.), or device related (processing power and space available in device, resolution, audio, battery, etc.). Human IF refers to the demographics of the consumer, such as the age, gender, their background, etc.

2.5 Development of the User Interface

As the demand for internet video-streaming increases, so does the competition between different video providers. Each video provider must ensure that their application is up to date with the latest technology and the latest features, and ensure that the user spends time on their application instead of switching to another app if they find it boring or frustrating to use their app. This means that, along with the architectural support, even the user interface (UI) needs to be optimized, since the user interacts with it.

The UI is built to be as attractive as possible to the user, be it the layout or design, or the arrangement of icons and buttons across the app. The application developer usually keeps track of the statistics and analytics of video demand from the general public, as well as user-specific preferences and provides a list of popular or recommended content, which the user can choose from. Artificial intelligence learning algorithms are used to assess the interests of the user, predict which type of videos the user might try to access and recommend more such content. Since all possible content from across the world is available to the user, it is necessary to ensure that the content is well-organized and presented to the user in the order of relevance, and analytics software comes to aid.

Since the primary purpose of streaming video over the internet is to ensure that the user can access the content at their convenience, it is necessary to incorporate features that facilitate this. This includes the provision to record content which is streaming or upcoming, downloading onto the device for offline viewing, marking videos for watching later, saving them, or marking them as favorites, etc. Even while streaming a particular video, there is a possibility to pause, skip few seconds, replay, stop and resume, or even seek to a desired time-point which is not possible in traditional cable TV.

To make the video assets attractive to the user, appropriate thumbnails are used to identify assets. Some applications make use of different thumbnails for different users based on the genre which the particular user is interested in, which helps to get more views. Thumbnails are also made to appear while the user seeks so that they can seek to a desired location on the timeline of the video asset. Other features that are included in video-streaming applications are the possibility to switch between languages of audio and subtitles. All these are stored in separate files, and can be requested and fetched from the server as per the user's request. Subtitles or closed-captioning (CC) are used to provide a translation of the audio that is being played in the form of text which appears on the screen. While traditionally these CC files are pre-written and stored, with technological advancements in artificial intelligence and speech recognition techniques, it is possible to analyze the words that are spoken and auto-generate the subtitle text in real-time for the user.

Generating revenue tactically is highly important for the video provider. Most video streaming applications come with a subscription fee which gives the user access to a certain set of channels and content. But there are also several applications that provide video streaming for free, and they utilize the system of ad-insertion. Though advertisements can be pre-attached to the content file and sent to the user initially itself, many video service providers make use of dynamic ad-insertion, in which ads are fetched in real-time when the user reaches a designated time-slot while playing the content. This is carried out by sending requests to an ad-server, which returns an ad of that particular time-slot duration, which is based on the user's interests, or the sponsors for the video, while ensuring that ads do not get repeated, since that badly affects user experience. This is also advantageous during live streaming, since the time for ad-slot is not predetermined.

3. SUMMARY

Video-streaming via the internet has become an integral part of our daily life and has even replaced traditional streaming over cable TV in many cases. Hence it requires a well-developed architecture which supports streaming effectively. In this paper, a review of some of the most important technological features required to stream video content over the internet is provided. CDNs enable global scalability of the content while helping in distribution of load across edge-servers. ABR helps streaming possible across a wide range of devices and accommodates for changing network bandwidth. Effective video compression techniques ensure that the size of the file is reduced to a minimum while maintaining the original quality. QoE acts as a better tool of measurement of the experience of the user, rather than just measuring the network performance characteristics. The User-Interface is designed to make the streaming experience comfortable for the user. All these features ensure that the experience of the user matches the expectations, while providing room for further development in this field.

REFERENCES

[1] J. Zhao, P. Liang, W. Liufu and Z. Fan, "Recent Developments in Content Delivery Network: A Survey", International Symposium on Parallel Architectures, Algorithms and Programming, 2019, pp. 98-106

[2] A. Binder, "Content Delivery Network Interconnect", Information Sciences & Technologies: Bulletin of the ACM Slovakia, 2017

[3] S. Saroju, K. P. Gummadi, R. J. Dunn, S. D. Gribble, H. M. Levy, "An Analysis of Internet Content Delivery Systems", 2002

[4] Y. Sani, A. Mauthe, C. Edwards, "Adaptive bitrate selection: A survey", IEEE Communications Surveys & Tutorials, 2017

[5] J. Kua, G. Armitage, P. Branch, "A survey of rate adaptation techniques for dynamic adaptive streaming over HTTP", IEEE Communications Surveys & Tutorials, 2017

[6] V. Sulochana, B. Shanthini, "A Survey on Adaptive Video Streaming Techniques with Cloud", International Journal of Pure and Applied Mathematics, 2018

[7] M. F. Hossain, M. Sarkar, S. H. Ahmed, "Quality of experience for video streaming: A contemporary survey", 13th International Wireless Communications and Mobile Computing Conference (IWCMC), 2017, pp. 80-84