EAI Endorsed Transactions
on Cloud Systems Research Article EAI.EU

A Novel and Optimal approach for Multimedia Cloud Storage and Delivery to reduce Total Cost of Ownership

Sreelakshmi Gollapudi[1,*], Kartikeshwar Rath[1], Tushar Vrind[1], Deepak TY[1] and Prasad R Rao[1]

[1]Samsung Semiconductors India R&D Centre, Samsung Electronics, Bangalore, India

Abstract

In this era of digital communication, and explosion of social media, users generate and share a lot of information most of which is audio visual content. This kind of multimedia content requires good amount of storage in the local device space as well at the network space. In the available parlance of multimedia cloud storage, when the content is streamed from the content server, the bit-stream is typically adapted depending on the available network bandwidth between the client and server session, for example by using Scalable Video Coding (SVC) technique. However, in case when the content is downloaded at the client for offline viewing, with say a resolution ‘Low-Res-1’, the multimedia clouds, do not offer additional mechanism to upgrade to a new resolution say ‘High-Res-2’, without downloading a new file version all over again. In this paper, we propose “MediaStratify” as a novel and optimal approach built on top of SVC to give a scalable solution for storing, sharing and upgrading the multimedia content for viewing offline. Based on the proposal, multimedia content will be stored as layers or ‘stratified’ and distributed over the cloud infrastructure. Through the devised protocol, the end node fetches the partial offsets (spatial, temporal or quality) and upgrades the files through reconstruction. Enterprise applications can utilize the scheme by installing the proposed novel combiner over the file transfer service, the solution can save network bandwidth and power consumption. The most important contribution is to bring down the Total Cost of Ownership (TCO) for any multimedia cloud or data center by reducing storage requirements by 50 ~ 74% over classical methods, yet achieve the goals of media hosting.

Keywords: Scalable Video Coding, Cloud, Data centre, Temporal Scalability, Spatial Scalability.

Received on 02 November 2019, accepted on 20 December 2019, published on 10 January 2020

doi: 10.4108/eai.5-11-2019.162596

1. Introduction

There has been explosive growth in the demand of various video-based applications – ranging from video telephony, video sharing, streaming and file sharing etc. ‘Always On’ mobility is the new normal. The moment a user goes online he goes on creating more content and data by using different application or services. By 2020 we would be creating the content of 44 Zeta Bytes (ZB) across the world – which means every user would be creating Mega Bytes’ (MB’) of data every second. This demands that we store data at some central location and the availability of it should be high as well as quick.

Usually any connected device end up in being connected to a data center on the cloud – which offers computing powers along with storage and networking support. In order to mitigate the never ending demands of application usage - ranging from handheld devices to smart devices, we need to bring more agility in compute, networking and storage nodes across the data centers.

Cloud technology has played a pivotal role in this as ‘Cloud is the new hardware’. It helped in the crucial part for nodes of the data center to be segregated. This gives cloud operators the freedom of implementing any strategy or use case (Infrastructure as Service (IasS) or Platform as Service (PasS) or Software as Service (SasS)) using the same infrastructure – on demand and without doing any physical changes as shown in Figure 1. IaaS provides virtualized computing resources over the internet. PaaS provides

[*]Corresponding author. Email: s.gollapudi@samsung.com

hardware and software tools needed for application development to users over the internet. SaaS is a software distribution model in which software is hosted in common place and users access it over the internet.

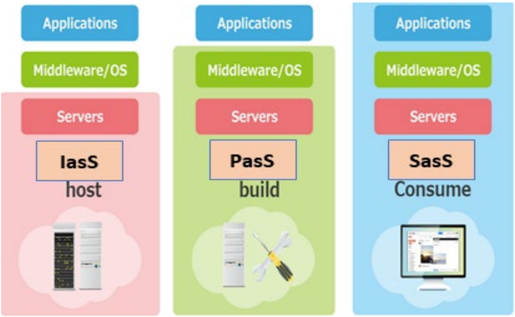

Figure 1. Deployment of Cloud Services

If we look at multimedia alone, is has resulted in generating more data on a regular basis and rapid developments in various network infrastructure nodes as well as more storage space requirements for a multimedia cloud server. While there are multimedia encoding schemes like Scalable Video Coding (SVC), which offers adaptive streaming of multimedia content depending on available network bandwidth between client and server, it has not been utilized in the distributed storage paradigm.

In case when the content is downloaded at the client for offline viewing, with say a resolution 'Low-Res-1', the multimedia clouds, do not offer additional mechanism to upgrade to a new resolutions say 'High-Res-2', without downloading a new file all over again. Use cases, where content is first downloaded and then used, for example, video analytics of surveillance content, or offline play back of multimedia content, complete download of a higher resolution content is not an attractive solution form network resources like storage, bandwidth and power perspective.

To solve the above mentioned problem, we propose a novel mechanism, which can split and store video content as layers in the cloud infrastructure and download on demand differential versions to enhance the content resolution at the client. From storage optimization perspective, it is further suggested to keep the differential version of the multimedia content at hybrid storage, to optimize the cost of the storage. The hybrid storage is usually an array of Hard Disk Drive (HDD) and Solid-State Disk Drive (SSD).

Through mathematical simulation derived from analytical model of multimedia cloud storage, we show that there is a decrease in Total Cost of Ownership (TCO) by over 50 to 74 % depending on the applied storage coding.

The rest of the paper is organized as follows. Section II covers the background on SVC. Section III discusses the proposed solution – "MediaStratify". Section IV covers the simulation model and results. Section V discusses the conclusions.

2. Approach – SVC Background

Before SVC came into picture Advanced Video Coding (AVC) [5] (simple H.264) was the standard method, one example of encoding non scalable video is shown in Fig. 2. I, intra frame is an independently coded frame and it can be decoded independently. P, prediction frame is coded with prediction from previous frames which means without the previous frame, decoding will not be successful. Due to poor radio conditions if one of the frame is lost, the recovery is not possible. To address the terminals with different spatial or frame rate or Signal to Noise Ratio (SNR) resolutions for the same video stream multiple versions has to be coded and saved in cloud server to stream for different applications. In simulcast single source streams to different destinations as shown in Figure 3. These destinations may have different spatial or frame rate or SNR resolutions.

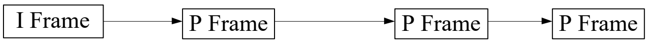

Figure 2. Advanced Video Coding (H.264)

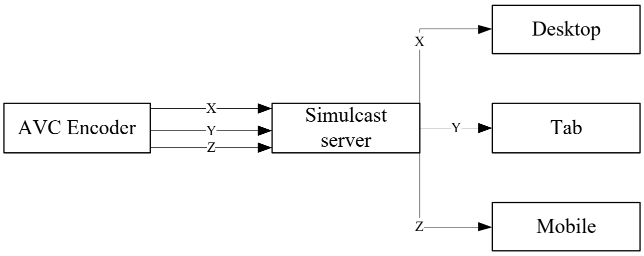

Figure 3. Simulcast Streaming

SVC is an extension of H.264 [1], SVC encodes the video signal as a set of layers. The various layers depend on each other, forming a hierarchy. A particular layer, together with the layers it depends upon, provides the information necessary to decode the video signal at a particular fidelity. Fidelity means one or more of spatial resolution, temporal resolution, or signal-to-noise ratio (SNR). The base layer, i.e., the layer that does not depend on any other layer, gives the lowest quality of the original video. Each enhancement layer once added improves the quality of the video in any one of the three dimensions (spatial, temporal, or SNR). SNR scalability provides the different video qualities for the same video stream maintaining the same temporal and spatial resolutions. In SNR scalability, the base layer encodes the coarsely quantized coefficients, transmits with moderate quality and lower bit rate. The difference between Non-quantized and coarsely quantized values will be finely quantized and encoded and transmitted in enhanced layer. Together with base layer enhancement layer provides the high SNR. Spatial scalability supports terminals with different resolutions. For example with just base layer Standard Definition TV (SDTV) can be supported and with adding enhanced layer it can support High Definition TV (HDTV). Temporal scalability supports terminals with different frame rates or temporal resolutions. As video is encoded into different layers, the layers which are important can be coded with high quality so even in poor radio conditions, the chances of recovery will be higher as base layer can be independently decodable. Similarly there is no need to create multiple streams to address the different spatial/temporal/SNR resolutions of terminals, as from one SVC encoded stream different terminal requirements can be met by choosing the appropriate layers according to terminal capability/network bandwidth or signal conditions. Typically, deployed in the streaming domain, SVC provides a network bandwidth aware scaling mechanism, where the user gets a better viewing experience [2]. SVC Encoded stream addressing different application requirements is shown below in Figure 4.

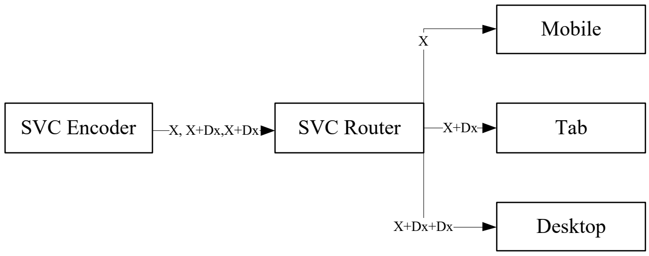

Figure 4. SVC Streaming

SVC covers two main parts - Video Coding Layer (VCL) and a Network Abstraction Layer (NAL) [1]. While the VCL creates a coded representation of the source content, the NAL formats these data and provides header information in a way that enables simple and effective customization of the use of VCL data for a broad variety of systems.

3. Proposed Solution: "Media Stratify"

SVC can generate temporal and spatial scalable encoded stream. The encoded stream is carried in various NAL units containing data and control information. In the below example given in Figure 5, there are three temporal layers. There is a base layer L0 (15 fps) along with two more enhancement layers L1 (7.5 fps) and L2 (7.5 fps). Though in the example only temporal scalability is shown for simplicity, spatial scalable layers can be defined for different resolutions. Spatial scalability can be combined with temporal (or SNR) scalability in completely independent way and different combinations such as 30fps720p, 30fps720p, 30fps1080p, 15fps1080p e.t.c can be created.

The proposed "MediaStratify" will split the encoded content (file containing NAL units of all layers) into different files depending upon the no of temporal layers. This is achieved with the help of 'MediaStratify.splitter' function. Each segregated file will contain all the NAL units corresponding to a specific temporal layer L0 or L1 or L2.While storing the segregated files, 'MediaStratify' will add special metadata to it. The metadata information would contain the encoding configuration (like fps, resolution, and bit rate) along with the 'media marker', which would help in managing the contents, useful in locating the required delta files to achieve the required scalability (like – increasing from 15fps to 30 fps). Additionally, another metadata file containing the indexing of the spitted files can be kept at the caching tier. Upon the request from the device or client application – for the video storage upgrade - the relevant enhancement file could be transferred from the storage units. The scenario is described in Figure 6 with more details.

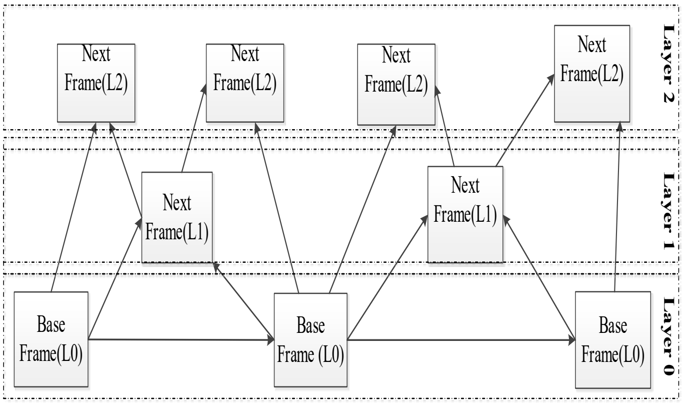

Figure 5. Temporal scalability [3]

As shown in Figure 6, the base file at the client is having 15 fps (L0) resolution. Upon request for "Fetch Upgrade" (for 22 fps) – the split file (having only the L1 Layer (7 fps) is located in the DC and the same is transferred over the network. At the client, 'MediaStratify.combiner' would parse through SVC control information present in the base file and the downloaded file, to combine the desired upgraded target file (22 fps in this case).

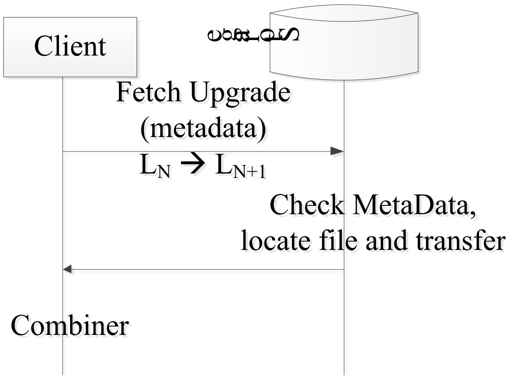
Client
Fetch Upgrade (metadata)
$L_N \rightarrow L_{N+1}$
Check MetaData, locate file and transfer
Combiner

Figure 6. Mechanism for upgrading content resolution

'MediaStratify' as a solution can be deployed as an enterprise application using the SaaS model of cloud. This would help to move out from a on premise setup to a scalable external cloud setup. The multimedia content while being stored on the cloud setup would be indexed as per the splitter logic and can be indexed to be stored across different

Table 1. File size description for a 90 minutes video file

	Size for Content Types (in TB) for media contents [6]				
Type of DC	4K	2K	1080p	720p	480p
Traditional DC	1066.95	442.39	260.23	104.09	26.02
"MediaStratify" - (temporal and spatial) DC	277.58	80.96	69.40	34.70	11.57
"MediaStratify" - (temporal) DC	474.20	196.62	115.66	46.26	11.57

storage nodes across the cloud. We can store the base files (L0) in all the caching tier of data centers. The other layers, can be stored depending on the user patterns and can be regionally cached.

4. Simulation Model and Evaluation Results

We adopted proposed scheme for a Multimedia Cloud Data Center (DC), which caters to the movie contents as captured in [4] and [6]. We simulated the storage requirement based on temporal and spatial versions, captured in Table 1 for uniform 90 minutes length of video files. L0 is 15 fps, L1 is 7.5 fps and L2 is 7.5 fps, so effectively “Fetch Upgrade” can scale up the file to 30 fps, by incremental fetches, while the traditional system would fetch the entire file. In the spatial scalable domain fetch upgrade can scale up to 4K. Based on the above Data Center Model, we simulated the storage requirements for the multimedia cloud with traditional design of keeping files with all resolutions and then with the proposed solution.

For the case of 90 minutes length of video file, when the Multimedia DC supports all versions till 4K video the effective DC size would be 469.37 TB, while with “MediaStratify” application to provide only temporal scalability, around 209 TB would suffice. When “MediaStratify” is applied to provide both temporal and spatial scalability in the content upgrade, only 122 TB would suffice. Table 2, captures data for all resolutions and gains are plotted in Figure 7.

Table 2. DC sizes with and without “MediaStratify”

Spatial Type	Total size (GB)	L0 (GB)	L1 (GB)	L2 (GB)
480p	1.3	0.65	0.32	0.32
720p	3.9	1.97	0.98	0.98
1080p	7.9	3.95	1.97	1.97
2K	9.2	4.6	2.3	2.3
4K	31.6	15.8	7.9	7.91

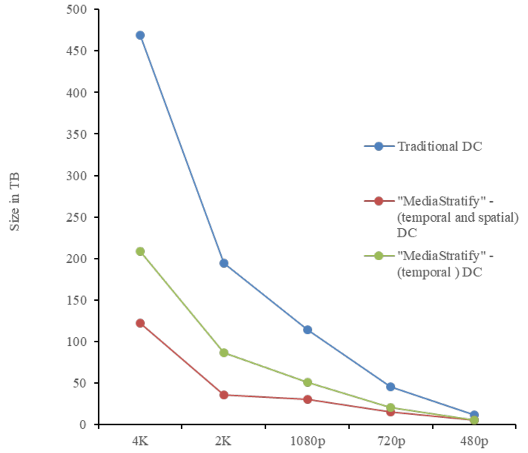

Figure 7. Storage savings due to "MediaStratify" for contents in [4]

As captured in Figure 7, we see an optimization of 74% in storage space at the data center. (((469.37 TB – 122 TB)/469.37 TB) * 100 = 74%), when both temporal and spatial encoding is applied to store layers of multimedia content. Fig. 8 shows the download time comparison of SVC vs “MediaStratify” for Ref [4].

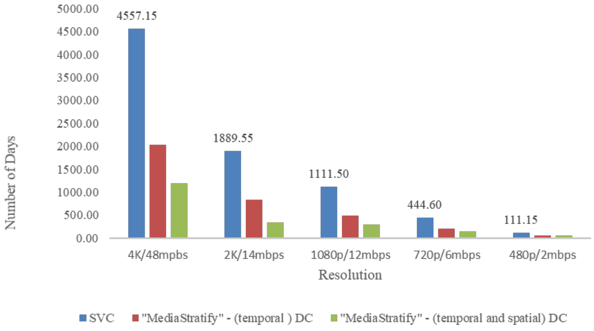

Figure 8. Download time comparison, "MediaStratify" vs SVC for contents in [4]

Another set of data [6] where the file length varying from 511 minutes to 7 minutes is used to simulate the storage requirements. Table 3, captures data for all resolutions for Ref [6] and storage gains are plotted in Figure 9.

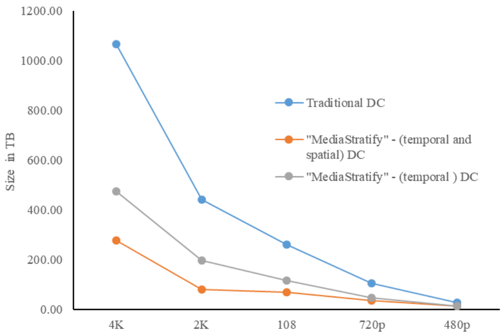

Figure 9. Storage savings due to "MediaStratify" for contents in [6]

The optimization is of 74% in storage space at the data center even for variable length files, when both temporal and spatial encoding is applied to store layers of multimedia content.

Figure 10 shows the download time comparison of SVC vs “MediaStratify” for Ref [6].

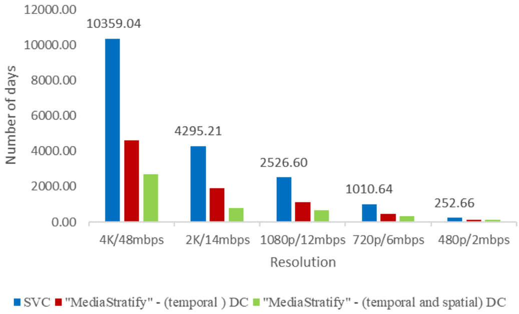

Figure 10. Download time comparison, "MediaStratify" vs SVC for contents in [6]

If only base layer is transmitted same amount of time is taken for downloading the file in SVC and “MediaStratify”, however if base layer needs to be upgraded to high quality more number of bits need to be transmitted in the case of SVC compared to “MediaStratify”. The results in Fig. 8 and Figure 10 shows 74% less time is needed with “MediaStratify” compared to SVC irrespective of length of the file and number of files. The comparison was done based on the assumption that file transfer is happening on 10mbps data rate wireless network without any retransmissions.

Table 3. DC sizes with and without “MediaStratify”

	(TB) Size of files for media contents [4]				
Type of DC	**4K**	**2K**	**1080p**	**720p**	**480p**
Traditional DC	469.37	194.62	114.48	45.79	11.45
Mediastratify DC (temporal & spatial scalable)	122.11	35.62	30.53	15.26	5.09
Mediastratify DC (temporal scalable)	208.61	86.50	50.88	20.35	5.09

5. Conclusion

By adopting “MediaStratify” there is a significant of savings in storage space of multimedia cloud DC, which can help reduce the TCO. This also lowers the storage requirements on the client side. In addition, a significant network bandwidth will be saved as only the differential content is transferred. Download time will be less compared to SVC. We propose the use of SasS model where we describe the functionality of both ‘MediaStratify.splitter’ and ‘MediaStratify.combiner’ for spatial scalability, temporal scalability and quality scalability, through which an effective storage saving is in the range of 50~ 74% depending on the applied encoding. The saving also translates to other aspects like system input outputs per second (IOPS), energy consumption – which is much beneficial for the DCs when annualized.

References

[1] Schwarz, Heiko etal. Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 17, NO. 9, SEPTEMBER (2007)

[2] Huysegems, Raf etal. SVC-based HTTP adaptive streaming, Bell Labs Technical Journal (Volume: 16, Issue: 4, (2012)

[3] Mathew, Manu, Application Report, SPRABG3 – November, Texas Instruments, http://www.ti.com/lit/an/sprabg3/sprabg3.pdf (2010)

[4] Stable benchmark dataset. 1 million ratings from 6000 users on 4000 movies, https://grouplens.org/datasets/movielens/1m/ (2003)

[5] http://info.vidyo.com/rs/vidyo/images/WP-Vidyo-SVC-Video-Communications.pdf

[6] https://www.kaggle.com/karrrimba/movie-metadatacsv