Impact of Distributed Caching on Video Streaming Quality in Information Centric Networks

Lingchao Kong, Jingyi Zhu, Rui Dai, Mohammad Nazmus Sadat
Department of Electrical Engineering and Computer Science
University of Cincinnati, Cincinnati, OH, USA
Email: konglo@mail.uc.edu, zhujy@mail.uc.edu, rui.dai@uc.edu, sadatms@mail.uc.edu

***Abstract*—The increasing demand of global video traffic presents a grand challenge for Internet service providers. This is particularly important for Smart Cities, where the citizens will be surrounded by various multimedia services. The Information Centric Networking (ICN) paradigm has been proposed to handle the vast amount of data, mobility of users, and provide better scalability and efficiency in terms of bandwidth demand. However, the distributed caching mechanism in ICNs brings new challenges in maintaining satisfactory Quality of Experience (QoE) for video streaming applications. Because of the caching capability of ICN routers, streaming from multiple sources will be common, and switching between content sources might affect QoE by inducing delays and consequently stalls in video playback. This paper investigates the impact of content source switching, which is brought by distributed in-network caching, on the QoE for video streaming in ICNs. Human subjective tests are conducted using videos generated from an ICN emulation platform, in which the videos are streamed to human users under different caching strategies and available bandwidths. Test results indicate that, as the number of source switches increases, the satisfaction with stalling decreases and thus resulting in lower QoE scores. The impact of stalling on the overall QoE is further quantified under different video clarity levels through several statistical methods.**

I. INTRODUCTION

With the ubiquity of smart mobile devices and the popularity of social media, the huge demand for video traffic presents a grand challenge for the current Internet infrastructure. Information-Centric Networking (ICN) [1] is proposed as a future Internet architecture to tackle the vast amount of data and provide better scalability and efficient content delivery in smart city applications. ICN identifies named content in a network and focuses on the data itself rather than the physical position of the data. ICN offers a number of benefits such as in-network caching, simpler configuration of network devices, native multicast support, and packet-level data security.

In-network caching allows faster content-fetching from nodes that are geographically closer to the user. By default, ICN follows the leave-copy-everywhere (LCE) caching strategy where popular contents are cached nearer to the end user. For a video streaming application, it is highly likely that the whole video will not be cached in all the nodes; rather, each node will cache a different amount depending on the popularity of the video among the end users connected to it. For example, 23% YouTube viewers watch only the initial 20% of a video [2]. This is also applicable to Netflix or Hulu, where the video sizes are even larger. Due to the limitation of cache size in ICN routers, it is not feasible to store additional video segments that are not being requested by users. Therefore, in contrast to traditional host-to-host video streaming applications in which a client downloads a video from a single source, a client in ICN may acquire the video from multiple sources because of distributed caching.

Switching between multiple content sources during video streaming could cause stalling, which is the momentary disruption in the fluidity of video playback [3]. Stalling is a major factor that degrades the Quality of Experience (QoE) at end users. The study on stalling events caused by network bottlenecks in [4] showed QoE is primarily influenced by the frequency and duration of stalling events. Moreover, stalling has worse impact on QoE than frame rate reduction [5]. For conventional networks, network congestion, collision, and packet loss are main contributing factors to stalling, and many techniques have been proposed to reduce stalling frequencies and lengths [6]–[8]. For ICNs, however, little is known about how source switching introduced by distributed caching could impact the QoE for video streaming. A few recent studies aimed at improving the QoE for video streaming in ICNs using strategies such as adaptive bitrate streaming [9], [10] and proactive content caching [11], but none of them studied the relationship between source switching and QoE in depth.

The goal of our research is to understand the impact of distributed caching and content source switching on the QoE for video streaming over ICNs. We consider the Content Centric Networks (CCN) [12] implementation of ICN in our work, which provides an emulation platform CCNx [13]. We conduct human subjective tests using videos generated from the emulation platform under different video content characteristics, file distribution strategies, and bandwidth conditions. Then we analyze the test results to answer two key questions: i) whether or not source switching could increase the frequency and the length of stallings in received videos; and ii) to what extent the stalling events contribute to the overall QoE scores. We also use several statistical methods to quantify the relationships between the overall QoE and the different quality-contributing factors.

The rest of this paper is organized as follows. The background and the experimental platform are introduced in Section II. The details of our human subjective test are described in Section III, and data analysis results are presented in Section IV, followed by conclusions in Section V.

978-1-5386-2937-6/17 $31.00 © 2017 IEEE
DOI 10.1109/ISM.2017.80

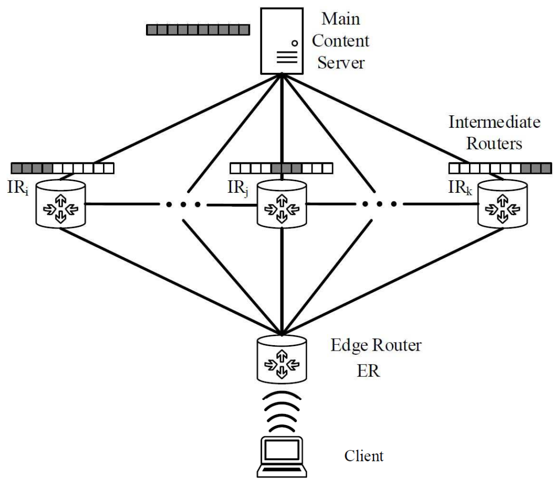


Fig. 1: Network topology

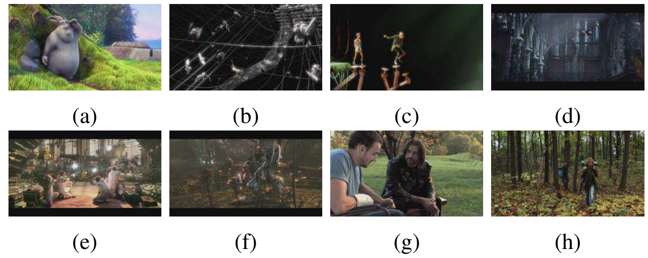


Fig. 2: Snapshots of video sequences. (a) Big Buck Bunny, (b)-(c) Elephants Dream, (d)-(f) Tears of Steel, (g)-(h) Valkaama

II. Background

Fig. 1 shows a typical ICN network topology and an example of content source switching during video streaming. A client is connected to an edge router ER and ER is in turn connected to the content server through several intermediate routers, such as IR_i, IR_j, and IR_k. To get a particular video, the client will send requests to ER. Depending on the availability of the content at ER, two scenarios may arise. First, the whole video may be cached at the Content Store (CS) of ER and thus ER will directly serve the video to the client. As discussed earlier, this is unlikely for most contents. Second, the video may be partially cached at ER or not cached at all. In this case, ER will obtain the complete or parts of video from the intermediate routers or the content server. In the example of Fig. 1, IR_i has cached the first 40% frames of a video. When IR_i finds that it does not have the rest of the video, it will switch to IR_j and IR_k, which can provide the middle 30% and the last 30% frames of the video. This results in a total number of two switches in the video streaming process. It should be noted that, even if the complete video is cached at one intermediate router, ER may still switch to another router with the same content if the link quality between ER and the first router deteriorates. If none of the routers contain a specific part of the video, the main content server will respond.

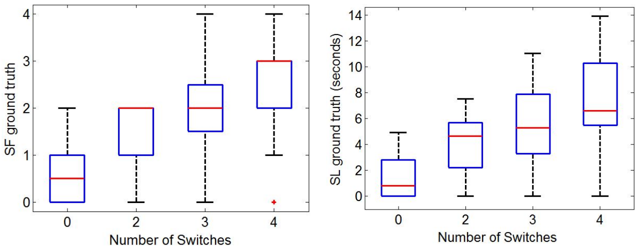
We installed CCNx 1.0 [13] (Distillery 2.0) in Ubuntu 14.04. All the routers and the main content server were connected through the Ethernet while the client and the edge router were connected through 802.11n wireless connection. The main content server had the complete video files and the intermediate routers contained different parts of the same video. Different cache distributions were used for different video files with up to 20% cache overlap between routers. This content distribution simulates the in-network caching of the CCN routers with the default LCE [12] caching strategy. LCE, also known as CCE (caching everything everywhere), is an on-path, homogeneous, and non-cooperative caching technique.

Each video file was chunked into 1200 byte segments and stored in the CS of intermediate routers. This segment size was chosen to prevent IP fragmentation of CCNx content object messages. Each intermediate router had a server application that maintained the videos in the CS and parsed the incoming requests to serve appropriate segments to the client. A modified version of the CCN-VLC plugin [14] was used for the streaming purpose on the client side. The requests sent from the client to the routers had the following format: *ccnx:/ccnx/fetch/Videoname.mp4/Chunknumber*. The client determines the chunk number using the file size information obtained from the server.

III. Description of the Subjective Test

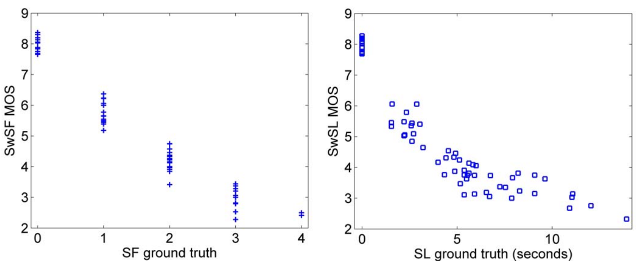
We conducted human subjective tests using videos generated from the aforementioned platform. We selected 8 raw video clips from four movies: *Big Buck Bunny* and *Elephants Dream* in Xiph.org Video Test Media Collection [15], *Tears of Steel* [16], and *Valkaama* [17]. These video clips contain different motion characteristics: half classified as high motion and half as low motion, based on the temporal information index defined by ITU-T P.910 [18]. The genres of the 8 videos cover animation, talk show, human with nature, science fiction and action movie, as shown in Fig. 2. Each video has 1280×720 pixels with a duration of 30 seconds. We encoded the eight raw videos in two bitrates (1500 kbps and 200 kbps) with the HM reference software (HEVC Standard) and generated 16 encoded videos. The choices of bitrates align with common bandwidth conditions in the network, and the two bitrates reflect different levels of clarity. Four file distribution cases for CCN caching were emulated: a video is distributed among 1, 3, 4, and 5 routers, resulting in 0, 2, 3, and 4 switches during the video streaming process. The 16 videos were transmitted over each of these cases, and finally 64 video samples were obtained. The channel bandwidth between the edge router and the client was set to 2600 kbps for higher quality videos and 350 kbps for lower quality ones. These bandwidth values were carefully selected to ensure that the channel had enough capacity to transmit videos without stalling in the absence of source switching.

In the subjective test, 36 participants (21 males, 15 females with normal vision aged from 18 to 34) rated the video samples in an indoor office with ordinary illumination. The methodologies for the assessment of video quality are based on



(a) Stalling Frequency (b) Stalling Length

Fig. 3: Ground truths of stalling frequency and stalling length



(a) Stalling Frequency (b) Stalling Length

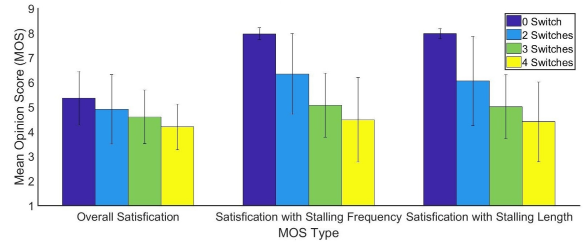
Fig. 4: Stalling ground truths vs. subjective scores

the recommendation of ITU-R BT.500-13, the single-stimulus method [19]. A graphical user interface was used to record individual ratings by playing 64 videos at random order. And the rating index with nine scales ranging from 1 (unbearable/poor) to 9 (satisfied/excellent) was used to evaluate the overall quality of the video i.e., overall mean opinion score (MOS), satisfaction with stalling length (SwSL), satisfaction with stalling frequency (SwSF), and video clarity, etc. The screening method recommended by BT.500-13 was used for screen the collected data [19]. The first four video ratings of each subject as ”dummy presentations” were removed for stabilizing opionions [19]. After examining the outliers [19], no observers were rejected.

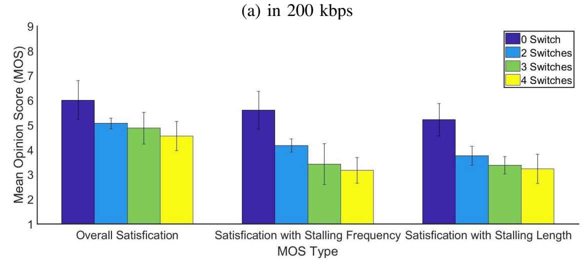
IV. Human Subjective Test Data Analysis

To understand the effect of source switching on the stallings in received videos, we first check the distributions of the ground truths of stalling frequency (SF) and stalling length (SL) under different number of switches (Fig. 3). The edges of the blue boxes are the 25th and 75th quartiles, and the central red marks are the medians. In Fig. 3 (a), the SF ground truth is defined as the times that stalling appear in each video; in Fig. 3 (b), the SL ground truth is defined as the total duration of each stalling event in a video. We find that as the number of switches increases, both the SF ground truth and the SL ground truth increase. An outlier is marked as a red dot in Fig. 3 (a). It corresponds to the video *TearsofSteel2* encoded under 200 kbps, where no stalling occurred even though there were four switches between the caching sources.

We then compare the subjective scores of SF and SL with their ground truths. Fig. 4 shows the relationships between the stalling ground truths of the 64 videos and their corresponding



(a) in 200 kbps



(b) in 1500 kbps

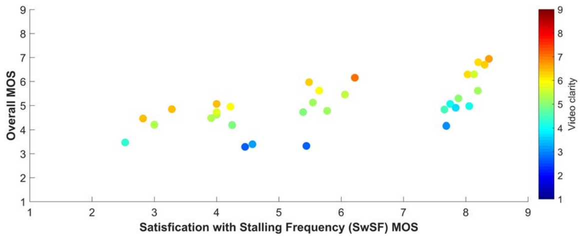
Fig. 5: Overall MOS, SwSF, and SwSL under four CCN conditions

subjective scores. In Fig. 4 (a), the SwSF MOS has negative correlation with the SF ground truth. In Fig. 4 (b), as the SL ground truth increases, the SwSF MOS decreases. These results indicate that the subjective evaluations of stalling align with the true stalling events in the videos.

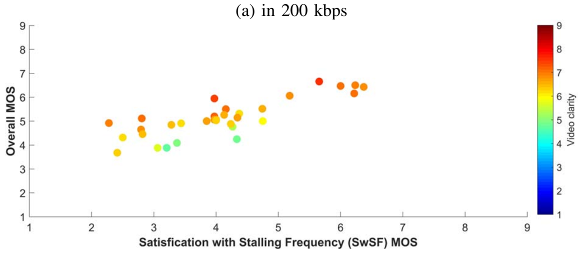
Fig. 5 presents the statistical results about the overall MOS, SwSF, and SwSL under different number of switches. The overall MOS, SwSF and SwSL all decline as the number of switches increases from 0 to 4. The overall MOS under 1500 kbps is better than the ones under 200 kbps, mostly due to the fact that higher bitrates correspond to higher clarity in the videos. On the other hand, the satisfaction scores with stalling (SwSF and SwSL) for videos encoded under 1500 kbps are lower than the ones encoded under 200 kbps, indicating that stallings are more likely to occur under higher bitrates.

We further analyze to what extent the stalling events could contribute to the overall QoE. Fig. 6 depicts the relationship between SwSF, video clarity and overall MOS in the two bitrates, in which the video clarity are represented by the color bars. For the high bitrate scenario, the video clarity MOSs are relatively high, and the overall MOS increases when SwSF increases. That is to say, stalling has greater impact on the overall MOS when the videos have high clarity levels. For the low bitrate scenario, the data are relatively dispersed, and there is no significant linear correlation between the overall MOS and SwSF. Furthermore, we analyze the correlation between the overall MOS and SwSF, SwSL, and video clarity for the two bitrates, and the results are shown in Table I. Similar conclusions could be drawn from this correlation analysis. In both the two bitrates, SwSF, SwSL are correlated with the overall MOS. Under 1500 kbps, SwSF and SwSL have more obvious linear correlation with the overall MOS, while under 200 kbps, video clarity also plays a significant role.

Finally, we use Principal Component Analysis (PCA)



(a) in 200 kbps



(b) in 1500 kbps

Fig. 6: Overall MOS vs. SwSF under different video clarity levels

Correlation Type	1500 kbps			200 kbps		
	SwSF	SwSL	Clarity	SwSF	SwSL	Clarity
Pearson	0.8279	0.8025	0.7043	0.6466	0.6551	0.8576
Kendall	0.5804	0.5131	0.5602	0.5576	0.5631	0.6276
Spearman	0.7467	0.6746	0.7502	0.7361	0.7539	0.8139

TABLE I: Correlation coefficient between Overall MOS with SwSF, SwSL and clarity

method to quantify the relationships between the different quality-contributing factors. The results are shown in Table II. The 1st and 2nd principal components (PC) maintain overwhelming superiority with 66.52% and 30.15%. Since cumulative contribution is larger than 95%, we can use the first two PCs to explain the total variability. The transform equation of PCs with original factors is $P_i = \sum_{j=1}^{j=5} c_{ij} \cdot X_j$, where P_i is the i principal component, c_{ij} is corresponding coefficient, X_j is the original factor. The five factors are movement consistency, video clarity, video quality constancy, SwSF, and SwSL, sequentially. The 1st PC transform vector is (0.0701, -0.1402, -0.0039, 0.7008, 0.6959), and the 2nd PC one is (0.4528, 0.7326, 0.5024, 0.0661, 0.0382). The transform vectors show that stalling satisfaction (SwSF and SwSL) dominate the 1st PC, and the video clarity contributes most to the 2nd PC.

V. Conclusion

In this paper, we have studied the impact of source switching, brought by distributed caching, on the QoE of video streaming in ICNs. We have found that when the source switching increases, both the frequency and length of stalling events increase, which was detected by human subjects. The overall satisfaction with the received videos declines when the number of switches increases. The results of correlation analysis and PCA indicate that the satisfaction with stalling is correlated with the overall MOS. Specifically, in high bitrate videos with high clarity, the satisfaction with stalling has obvious linear correlation with the overall MOS; while in low bitrates with lower video clarity, both the satisfaction with stalling and the clarity play significant roles in the overall MOS.

Principal Component	1st	2nd	3rd	4th	5th
Eigenvalue	6.5327	2.9608	0.1883	0.0857	0.0529
Proportion (%)	66.52	30.15	1.92	0.87	0.54
Cumulative (%)	66.52	96.67	98.59	99.46	100.00

TABLE II: The results of principal component analysis

References

[1] G. Xylomenos, C. N. Ververidis, V. A. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. V. Katsaros, and G. C. Polyzos, "A survey of information-centric networking research," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 1024–1049, 2014.

[2] D. K. Krishnappa, D. Bhat, and M. Zink, "Dashing youtube: An analysis of using DASH in youtube video service," in *Local Computer Networks (LCN), 2013 IEEE 38th Conference on*. IEEE, 2013, pp. 407–415.

[3] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hobfeld, and P. Tran-Gia, "A survey on quality of experience of http adaptive streaming," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 469–492, 2015.

[4] T. Hoßfeld, M. Seufert, M. Hirth, T. Zinner, P. Tran-Gia, and R. Schatz, "Quantification of youtube QoE via crowdsourcing," in *Multimedia (ISM), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 494–499.

[5] Q. Huynh-Thu and M. Ghanbari, "Temporal aspect of perceived quality in mobile video broadcasting," *IEEE Transactions on Broadcasting*, vol. 54, no. 3, pp. 641–651, 2008.

[6] F. Pervez and M. S. Raheel, "QoE-based network-centric resource allocation for on-demand uplink adaptive HTTP streaming over LTE network," in *Application of Information and Communication Technologies (AICT), 2014 IEEE 8th International Conference on*. IEEE, 2014, pp. 1–5.

[7] K. Miller, E. Quacchio, G. Gennari, and A. Wolisz, "Adaptation algorithm for adaptive streaming over HTTP," in *Packet Video Workshop (PV), 2012 19th International*. IEEE, 2012, pp. 173–178.

[8] W. U. Rahman, D. Yun, and K. Chung, "A client side buffer management algorithm to improve QoE," *IEEE Transactions on Consumer Electronics*, vol. 62, no. 4, pp. 371–379, 2016.

[9] S. Awiphan, T. Muto, Y. Wang, Z. Su, and J. Katto, "Video streaming over content centric networking: experimental studies on planetlab," in *Computing, Communications and IT Applications Conference (ComComAp), 2013*. IEEE, 2013, pp. 19–24.

[10] Z. Liu and Y. Wei, "Hop-by-hop adaptive video streaming in content centric network," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–7.

[11] K. Kanai, T. Muto, J. Katto, S. Yamamura, T. Furutono, T. Saito, H. Mikami, K. Kusachi, T. Tsuda, W. Kameyama *et al.*, "Proactive content caching for mobile video utilizing transportation systems and evaluation through field experiments," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 8, pp. 2102–2114, 2016.

[12] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*. ACM, 2009, pp. 1–12.

[13] "CCNx project web site," http://www.ccnx.org, accessed: 2017-06-10.

[14] "CCN-VLC-AccessModule," https://github.com/PARC/CCN-VLC-AccessModule, accessed: 2017-06-10.

[15] "Xiph.org Video Test Media Collection," https://media.xiph.org/video/derf/, accessed: 2017-07-31.

[16] "Tears of Steel — Mango Open Movie Project," https://mango.blender.org/, accessed: 2017-07-31.

[17] "Valkaama," http://www.valkaama.com/, accessed: 2017-07-31.

[18] I. Recommendation, "P. 910,subjective video quality assessment methods for multimedia applications,," *International Telecommunication Union, Tech. Rep*, 2008.

[19] ——, "Bt. 500-13,methodology for the subjective assessment of the quality of television pictures,," *International Telecommunication Union, Tech. Rep*, 2012.