Peer-Assisted Content Distribution Aided by Video Popularity Evolution Model

Mohamad Dikshie Fauzie* Achmad Husni Thamrin* Jun Murai† *Graduate School of Media and Governance †Faculty of Environment and Information Studies
Keio University, 252-0882 Kanagawa, Japan
dikshie@sfc.wide.ad.jp husni@ai3.net jun@wide.ad.jp

***Abstract*—Generally content distribution network (CDN) have adopted two different architectures: client-server model that's is the most common architecture model and peer-assisted CDN. In client-server model, clients download content dedicated and geographically managed servers while in peer-assisted model, clients download content from each other client.**

***Index Terms*—P2P, CDN**

I. Introduction

Streaming content, especially video, represents a significant fraction of the traffic volume on the Internet, and it has become a standard practice to deliver this type of content using Content Delivery Networks (CDNs) such as Akamai and Limelight for better scaling and quality of experience for the end users. For example, YouTube uses Google cache and MTV uses Akamai in their operations.

With the spread of broadband Internet access at a reasonable flat monthly rate, users are connected to the Internet 24 hours a day and they can download and share multimedia content. P2P (peer to peer) applications are also widely deployed. In China, P2P is very popular; we see many P2P applications from China such as PPLive, PPStream, UUSe, Xunlei, etc. [1]. Some news broadcasters also rely on P2P technology to deliver popular live events. For example, CNN uses the Octoshape [2] solution that enables their broadcast to scale and offer good video quality as the number of users increases.

From the Internet provider point of view, the presence of so many always-on users suggests that it is possible to delegate a portion of computing, storage and networking tasks to the users, thus creating P2P networks where users can share files and multimedia content. Starting from file sharing protocols, P2P architectures have evolved toward video on demand and support for live events.

Alternatively, video contents can be efficiently distributed on services offered by managed network architectures and CDN companies. The major issues of CDN are high deployment cost and good but not unlimited scalability in the long term. Given the complementary features of P2P and CDN, in recent years some hybrid solutions have been proposed and applied to the operational of CDN [3]–[5] to take the best of both approaches. In Peer assisted CDN, users can download content from CDN nodes from or other users or peers. A user may cache the content after download to serve requests from other users. Due to the complexity of the behavior of peers, the process should be done in the home gateway user where the ISP can control it.

In this work, we will revisit Guo et al., [6] work's PROP as basis to evaluate of the peer-assisted CDN and propose an improvement the model for the PROP. We also examine the characteristics of Internet VoD by investigating real-world data sets obtained from Youtube. In P2P assisted CDN for video on demand (VoD), most of researcher assume that catalog of video popularity rank is already established following zipf distribution. This become basis for P2P assisted CDN model in PROP [6]. Our work is quite different whereas we will use VoD view popularity to aid the PROP model. We use Youtube VoD view model for this purpose. A twofold of our contributions as follows: (1) We use the idea of VoD view popularity model to aid the PROP model. To our knowledge, the combination of PROP model and VoD view popularity model is the first. (2) From simulation-based experiments, we find that peer contributions become higher than the PROP model.

Our paper presentation as follows: (1) we describe related work in sect.II; (2) we explain detail of Youtube popularity evolution model in sect.III; (3) we explain the caching strategy for CDN and peer in sect.IV; (4) we explain our simulation design, simulator, and its evaluation in sect.V. Finally, we present our conclusions in section VI.

II. Related Work

Content Distribution Networks with peer assist have been successfully deployed on the Internet, such as Akamai [8], [3] and LiveSky [9]. The authors of [8] examine the risks and benefits of peer-assisted content distribution in Akamai and measure the effectiveness of its peer-assisted approach. The authors of [3] conclude from two real world traces that hybrid CDN-P2P can significantly reduce the cost of content distribution and can scale to cope with the exponential growth of Internet video content. Yin et al. [9] described commercial operation of a peer-assisted CDN in China. LiveSky solved several challenges in the system design, such as dynamic resource scaling of P2P, low startup latency, ease of P2P integration with the existing CDN infrastructure, and network friendliness and upload fairness in the P2P operation. Xu et al. [10] using game theory, showed that the right cooperative profit distribution of P2P can help the ISP to maximize the utility. Their model can easily be implemented in the context of current Internet economic settlements. Misra et al. [11] also mentioned the importance of P2P architecture to support content delivery networks. The authors use cooperative game theory to formulate simple compensation rules for users who run P2P to support content delivery networks.

The idea of telco- or ISP-managed CDN has been proposed in recent years. The complexity of the CDN business encourage telcos and ISPs to manage their own CDN, rather than allow others to run CDNs on their networks. It has been shown that it is cost effective [12] [13]. Kamiyama et al. [14] proposed optimally ISP operated CDN. Kamiyama et al. mentioned that, in order to deliver large and rich Internet content to users, ISPs need to put their CDNs in data centers. The locations are limited while the storage is large, making this solution effective, using optimum placement algorithm based on real ISP network topologies. The authors found that inserting a CDN into an ISP's ladder-type network is effective in reducing the hop count, thus reduce total link cost. Cisco has initiated an effort to connect telco- or ISP-managed CDNs to each other, to form a CDN federation [12] using open standards [15]. They argue that the current CDN architecture is not close enough to the users and ISPs can fill this position.

The idea of utilizing the user's computation power to support ISP operation is not new. The Figaro project [16] proposed residential gateway as an integrator of different networks and services, becoming an Internet-wide distributed content management for a proposed future Internet architecture [16]. Cha et al., [17] performed trace analysis and found that an IPTV architecture powered by P2P can handle a much larger number of channels, with limited demand for infrastructure compare to IP multicast. Jiang et al. [18] proposed scalable and adaptive content replication and request routing for CDN servers located in users' home gateways. Maki et al. [19] propose traffic engineering for peer-assisted CDN to control the behavior of clients, and present a solution for optimizing the selection of content files. Mathieu et al., [20] are using data gathered from France telecom network to calculate reduction of network load if customers are employed as peer-assisted content delivery.

Guo et al., [6] is closest with our work because we use that work as comparison and we use author's utility function. The author proposed local system (local counter) to calculate the segment popularity in peer-assisted proxy system. The author use popularity for proxy cache replacement strategy. In peer side, the author use utility function for cache replacement strategy. The utility function is also function from popularity. While the authors successfully show that the results are very good, the peer-assisted system behavior over time is not explain because the author focus on local properties such as proxy cache size variations and peer cache size variations. The explanation of the optimal number of replicas is not also clear because unavailable information when the snapshot is taken. In our work, we complement Guo et al., [6] work with VoD viewing popularity evolution model and describe the behavior of the peer-assisted CDN over the time.

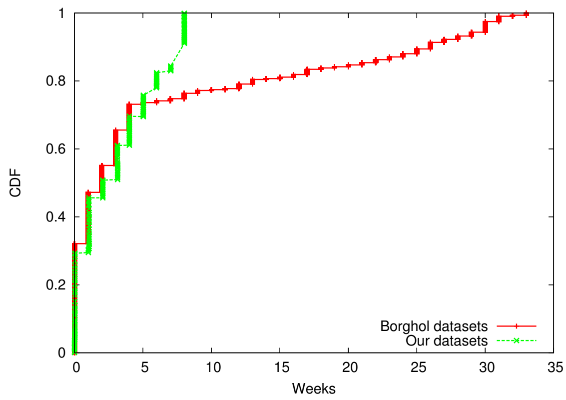


Fig. 1: Time to peak empirical distribution.

III. Characterizing Internet VoD Popularity

Before analyzing the system description and video caching, we first examine the popularity characteristics of Internet VoD services. We use YouTube as example of VoD service. The studies of content popularity evolution are mostly considered in short time periods. Borghol et al., [7] measure the evolution of content popularity in long periods (36 weeks) in which view count statistics of Youtube. Complement with Borghol et al., [7] work, we also did same measurement as Borghol et al., [7] for eight weeks from October-November 2013. We collected number of views, upload time YouTube videos on a weekly basis. We use YouTube's API to sampling popular videos. The API provides a call that returns details on 20 popular videos. Finally, we combine our datasets with Borghol et al., [7] datasets.

In datasets, we have one-week spacing between consecutive snapshots. We can get how many times the video was view during the one-week period since last week or since snapshot $(i-1)$. As same Borghol et al., [7] work, we define time-to-peak for a video as its age (time since upload) at which its weekly viewing rate is the highest during measurement (from the first week until end of measurement). The time-to-peak distributions is shown in fig.1. Figure 1 shows that around three-quarters of a large fraction videos peak within the first six weeks since their upload and beyond six weeks we have uniform distribution thus the time-to-peak is exponentially distributed mixture with uniform distribution. Our finding coherence with Borghol et al., [7]. To estimate the the rate parameter of exponential part of time-to-peak distribution, we use Maximum Likelihood Estimation (MLE) [21]. Using MLE method, we can get exponential parameter $\lambda = 0.61$. For weekly views distribution, Borghol et al., [7] found that beta distribution is a good model to explain video views evolution thus we follow Borghol et al., [7] for weekly views distribution model.

IV. System Description

In this paper, we consider a peer-assisted CDN system. There are two main components: (1) the CDN servers which at a minimum consist content delivery platform and control plane platform. (2) Clients which request and downloads the videos. In addition, clients form a self-organized P2P overlay network.

The CDN servers are maintained by a CDN company or a content provider company (e.q Netflix) or an Internet service provider (ISP). In peer-assisted CDN, a peer caches the videos that it has downloaded. Peers independently manage their cache localy. When a peer join the system video replica is cached. When a peer leaves the system video replica is evicted. In both process, a peer always reports to CDN thus CDN knows the status of a peer.

A. *Peer caching strategy*

In a peer side, we follow Guo et al., [6] for cache replacement strategy. We define the utility function for peer replacement as follows:

$$u = \frac{(f(p) - f(p_{min})) \times (f(p_{max}) - f(p))}{r^{\alpha+\beta}} \quad (1)$$

p represents popularity of the video, p_{min} represents estimation of minimum popularity in P2P system, p_{max} represents estimation of maximum popularity in P2P system, r represents the number of replicas of the video in the system, and $f(p)$ is monotonic non-decreasing function. α and β are the adjustment factor.

The CDN can calculate p_{min} and p_{max} then propagate to the P2P system. We choose the video with the smallest u value as the candidate to be replaced when a peer's cache capacity is full. In the next section (sect.V), we will show how popularity evolution knowledge is used to simplified our calculation for utility function.

B. *Peer caching strategy*

Equation 1 express the utility function used by peer for cache replacement. Since we can estimate before-peak week, at-peak week, and after-peak week of video, we modified the utility function as follows:

- In before-peak and after-peak phases, we assume that requests to video are low thus we only consider minimum popularity.
$$u = \frac{f(p) - f(p_{min})}{r^{\alpha+\beta}}. \quad (2)$$
- In at-peak phase, we assume that request to video are high thus we only consider maximum popularity.
$$u = \frac{f(p_{max}) - f(p)}{r^{\alpha+\beta}}. \quad (3)$$

V. Evaluation

In order to evaluate the proposed cache strategy using before-peak, at-peak, and after-peak information from VoD model, we have to compare our model to PROP model [6].

A. *Simulation Design*

An event driven simulator is developed using Python for this purpose and we use Youtube VoD model as video catalog of the simulated Internet VoD system in our experiment. In our simulator, time is divided into rounds. During a round, a peer request a video.

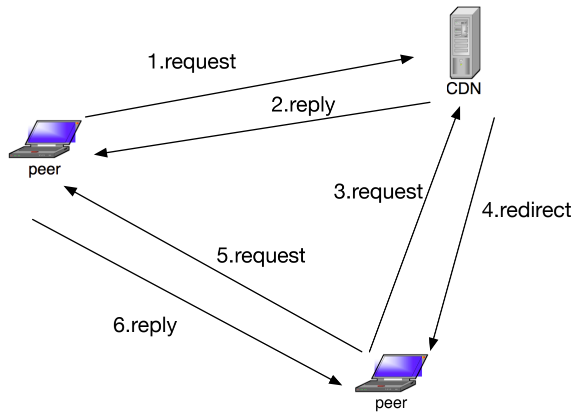


Fig. 2: Peer interaction in simulator.

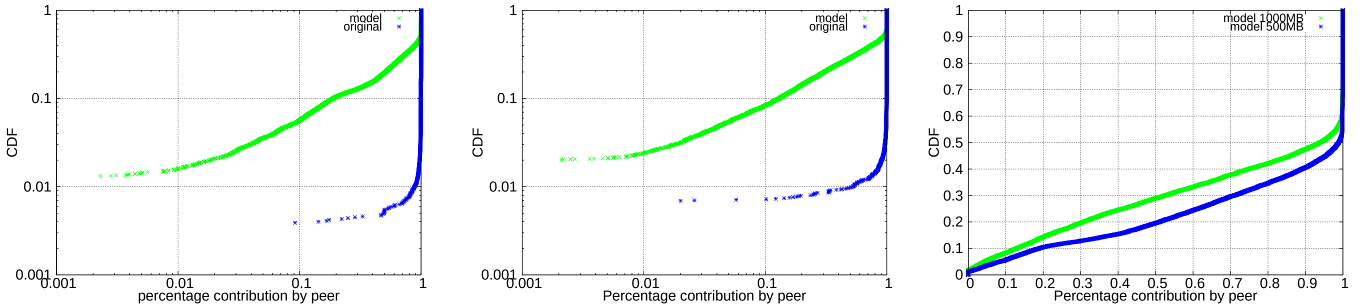
B. *Peer and CDN server interaction*

In fig.2, we describe the process of a peer that requests a video in simulator. peer and cdn are implemented in object oriented model. When a peer requests a video, it always goes to a CDN server (step 1). The CDN provides the videos to the peer (step 2). If there is another peer request same video, that request will go to CDN (step 3). CDN will check its record to see if there are some peers cache that requested video. If there are some peers cache that requested video, CDN will reply with redirect message that asking a peer to download requested video from other peer (step 4). If there are no peers have requested video, CDN will serve the video. A peer then can request the video to other peer and get the video (step 5 and step 6). From this description, we can see that deploying peer-assisted CDN can save some traffic since the clients which form P2P network can sharing the contents or videos.

1) Catalog Generator: In catalog generator, we assume peer request a video to CDN following poisson process with a mean rate $\lambda = 1.1$ [22] and we made it 3600 videos per hour, finally we generate video request for 360 days of simulation thus we have 31104000 requests by peers. How a peer choice a video, we will explain in next paragraph.

First of all, we calculate the number of videos at-peak time as follows: sample N value from the time-to-peak distribution and determine the number of videos n_j^{at} that peak at week j. Total number of video $N = n_j^{before} + n_j^{at} + n_j^{after}$.
Next, we determine view count terminus which the number of final view count of video. In view count terminus, we assume that a video will not get big additional view after at-peak phase. We assign view count terminus randomly from datasets. After determining view count terminus, we assign beta distribution parameter for every video.

Since we can estimate the time of at-peak phase for each video, we know the mode of beta distribution value and we can calculate α and β value using the mode of distribution formula: $m = \frac{\alpha-1}{\alpha+\beta-2}$. We assign α value randomly between 1 and 2 thus we can calculate β value. With the knowledge of beta distribution of every video and its view count terminus, we can know the view count and view rate of every video as function from time. The knowledge of view count and view

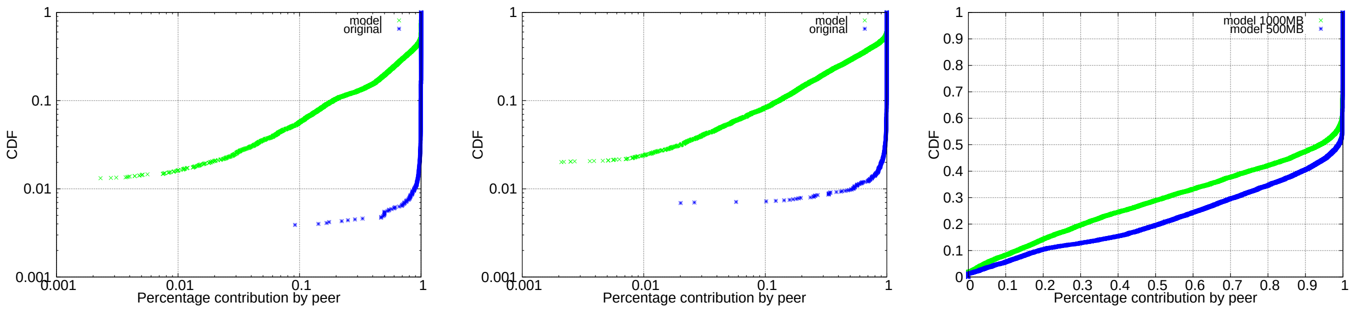


(a) CDF percentage of contribution of peers where peer capacity 500MB and CDN capacity 10GB. x and y axis in logscale.

(b) CDF percentage of contribution of peers where peer capacity 1000MB and CDN capacity 10GB. x and y axis in logscale.

(c) CDF peer contribution percentage between peer with 500MB capacity and 1000MB capacity and CDN capacity 10GB.

Fig. 3: CDF peer contribution for peer capacity 500MB and 1000MB, CDN capacity 10000MB. Model refers to our work and original refer to PROP [6].

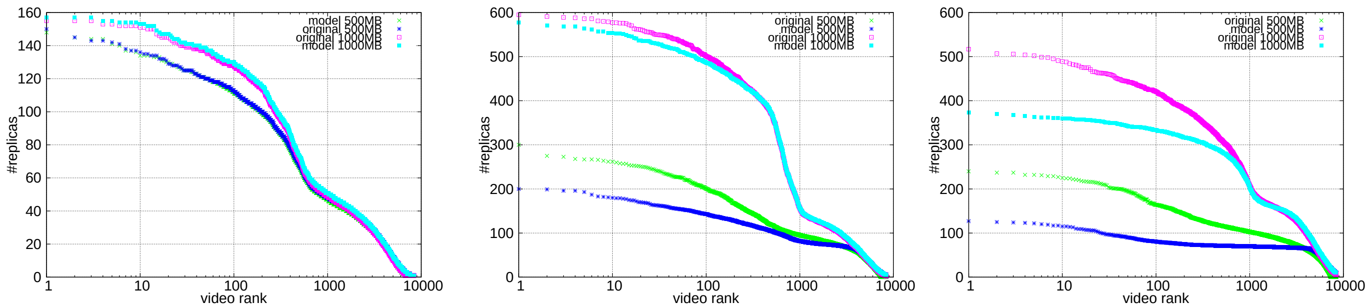


(a) CDF percentage of contribution of peers where peer capacity 500MB and CDN capacity 20GB. x and y axis in logscale.

(b) CDF percentage of contribution of peers where peer capacity 1000MB and CDN capcity 20GB. x and y axis in logscale.

(c) CDF peer contribution percentage between peer with 500MB capacity and 1000MB capacity and CDN capacity 20GB.

Fig. 4: CDF peer contribution for peer capacity 500MB and 1000MB, CDN capacity 20000MB. Model refers to our work and original refer to PROP [6].



(a) Number of replicas by popularity at $t = 10$day for peer capacity 500MB.

(b) Number of replicas by popularity at $t = 6$week for peer capacity 1000MB.

(c) Number of replicas by popularity at $t = 10$week for peer capacity 1000MB.

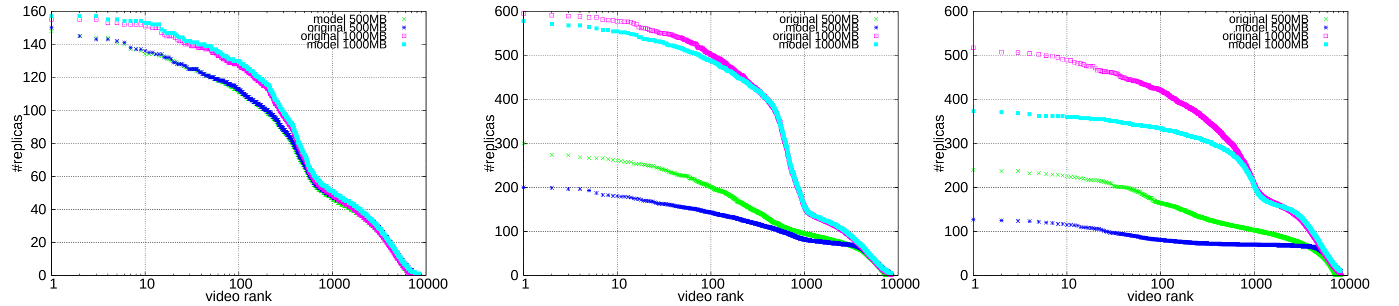
Fig. 5: Distribution number of video replicas at snapshot $t = 10$day, $t = 6$week, and $t = 10$week for peer capacity 500MB and 1000MB, CDN capacity 10GB.

rate, will be used top generate a video choice.

For video choice, we estimate that a peer will choose video proportionally considering view count and view rate of the video. We can get view count and view rate from probability distribution function (PDF) and cumulative distribution function (CDF) of beta distribution above multiply by video's view count terminus. In last step, we assign file size of video randomly between 1MB and 200MB. Finally, we have a catalog that consists of: video to be chosen, time when uploaded, view count terminus, at-peak week, and video size.

2) Simulation Parameters: The simulation parameters are follows:

- Length: 360 days.



(a) Number of replicas by popularity at t = 10day for peer capacity 500MB and 1000MB, CDN capacity 20GB.

(b) Number of replicas by popularity at t = 6week for peer capacity 500MB and 1000MB, CDN capacity 20GB.

(c) Number of replicas by popularity at t = 10week for peer capacity 500MB and 1000MB., CDN capacity 20GB.

Fig. 6: Distribution number of replicas at snapshot t = 10day, t = 6week, and t = 10week for peer capacity 500MB and 1000MB, CDN capacity 20GB.

- Video size: random between 1MB and 200MB.
- Peer capacity: [500MB,1000MB].
- CDN capacity: 10000MB and 20000MB.
- Number of peers: 100000.
- Number of videos: 10000.

We compare our proposed improvement of PROP to original PROP [6] implementation.

C. *Result and Discussion*

Figure 3 shows CDF peer contribution for our model compare to PROP for peer capacity 500MB and 1000MB, with CDN capacity 20GB. One dot in a figure means how many percentage a video delivered by peers during simulation.z Moreover fig. 3a shows the comparison of CDF of peer contributions between our model and PROP for peer capacity 500MB. Figure 3b shows the comparison of CDF of peer contribution between our model and PROP for peer capacity 1000MB. Figure 3c we compare our model between peer capacity 500MB and peer capacity 1000MB. Peer capacity 1000MB gives higher peer contribution because additional space makes a peer can cache more videos. From fig. 3b and fig. 3c, We can see that our model can gain higher peer contributions than PROP. We do significance statistical testing if our model has significantly different from PROP. We use Kolmogorov-Smirnov (KS) test for this purpose. KS-test tries to determine if two datasets differ significantly. It has the advantage of making no assumption about the distribution of data. KS-test reject the null hypothesis of no difference between datasets if p-value is small. For peer capacity 500MB when we compare our model and PROP, we get p-value $0.5e-005$. In peer capacity 1000MB case, we get p-value $0.46e-005$. Because both p-values are below 1%, we can reject null hypothesis that both data are the same thus our results are significant.

Figure 4 shows CDF peer contribution for our model compare to PROP for peer capacity 500MB and 1000MB, with CDN capacity 20GB. Figure 4a shows the comparison of CDF of peer contributions between our model and PROP for peer capacity 500MB. Figure 4b shows the comparison of CDF of peer contributions between our model and PROP for peer capacity 1000MB. From fig. 4 and fig. 4a show that our model can gain higher peer contributions than PROP. We also do significance statistical testing for both cases. We get p-value $0.5e-005$ for peer capacity 500MB case and p-value $0.47e-005$ for peer capacity 1000MB case. Because the p-values less than 1%, therefore the results are signifincant. Finally, fig. 4c shows the comparison between peer capacity 500MB and peer capacity 1000MB in our model. This figure shows that peer capacity 1000MB gives more contributions than peer capacity 500MB because additional space in peer makes a peer can cache more videos.

Figure 5 shows distribution of the number of replicas at snapshot t = 10day, t = 6week, and t = 10week ranked by video popularity for peer capacity 500MB and 1000MB, CDN capacity 10GB between our model and PROP. Figure 5a shows distribution of the number of replicas comparison for peer capacity 500MB and 1000MB and CDN capacity 10GB sorted by popularity rank at snapshot t = 10day. This figure shows that on snapshot t = 10day there are no many different for peer capacity 500MB between our model and PROP except for peer capacity 1000MB PROP has lower number of replicas than our model for popular video rank between 1 until 1000 beyond that the number of replicas is same. Figure 5b shows distribution of the number of replicas comparison for peer capacity 500MB and 1000MB and CDN capacity sorted by popularity rank at snapshot t = 6week. Both peer capacity 500MB and 1000MB show that our model has lower number of replicas compare to PROP for popular video rank between 1 and 1000 for peer capacity 500MB while for peer capacity 1000MB our model gives lower replicas for popular video rank between 1 and 400. Peer capacity 1000MB gives higher replicas than 500MB because additional capacity make more room for peer to cache videos. Figure 5c shows distribution of the number of replicas comparison for peer capacity 500MB and 1000MB and CDN capacity 10GB sorted by popularity rank at snapshot t = 10week. Again, our model gives lower number of replicas compare to PROP. Snapshot t = 10 weeks gives lower number of replicas than snapshot at t = 6week because in our model the utility function knows that t = 10week is at after-peak phase thus the model only considering minimum popularity for all videos.

VI. Conclusion and Future Work

This paper presents a scheme for a ISP managed peer-assisted CDN model that Some areas of improvement that we have identified for future are: We are also very interested to include energy trade off this peer-assisted CDN architecture in order to know how much energy saving by ISP and how much increase of energy at users home gateway side in this architecture.

Acknowledgment

The authors would like to thank Internet research laboratory member at Keio University and anonymous reviewer.

References

[1] L. Vu, I. Gupta, K. Nahrstedt, and J. Liang, "Understanding overlay characteristics of a large-scale peer-to-peer iptv system," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 6, no. 4, pp. 31:1–31:24, Nov. 2010. [Online]. Available: http://doi.acm.org/10.1145/1865106.1865115

[2] Octoshape, "Octoshape," http://www.octoshape.com/cnn-com-using-octoshapes-p2p-for-live-feed/.

[3] C. Huang, A. Wang, J. Li, and K. W. Ross, "Understanding hybrid cdn-p2p: why limelight needs its own red swoosh," in *Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, ser. NOSSDAV '08. New York, NY, USA: ACM, 2008, pp. 75–80. [Online]. Available: http://doi.acm.org/10.1145/1496046.1496064

[4] H. Jiang, J. Li, Z. Li, and J. Liu, "Efficient hierarchical content distribution using p2p technology," in *Networks, 2008. ICON 2008. 16th IEEE International Conference on*, dec. 2008, pp. 1 –6.

[5] H. Yin, X. Liu, T. Zhan, V. Sekar, F. Qiu, C. Lin, H. Zhang, and B. Li, "Design and deployment of a hybrid cdn-p2p system for live video streaming: experiences with livesky," in *Proceedings of the 17th ACM international conference on Multimedia*, ser. MM '09. New York, NY, USA: ACM, 2009, pp. 25–34. [Online]. Available: http://doi.acm.org/10.1145/1631272.1631279

[6] L. Guo, S. Chen, and X. Zhang, "Design and evaluation of a scalable and reliable p2p assisted proxy for on-demand streaming media delivery," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 18, no. 5, pp. 669–682, May 2006.

[7] M. Zhao, P. Aditya, A. Chen, Y. Lin, A. Haeberlen, P. Druschel, B. Maggs, B. Wishon, and M. Ponec, "Peer-assisted content distribution in akamai netsession," in *Proceedings of the 2013 Conference on Internet Measurement Conference*, ser. IMC '13. New York, NY, USA: ACM, 2013, pp. 31–42. [Online]. Available: http://doi.acm.org/10.1145/2504730.2504752

[8] H. Yin, X. Liu, T. Zhan, V. Sekar, F. Qiu, C. Lin, H. Zhang, and B. Li, "Livesky: Enhancing cdn with p2p," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 6, no. 3, pp. 16:1–16:19, Aug. 2010. [Online]. Available: http://doi.acm.org/10.1145/1823746.1823750

[9] K. Xu, Y. Zhong, and H. He, "Can p2p technology benefit eyeball isps? a cooperative profit distribution answer," *CoRR*, vol. abs/1212.4915, 2012.

[10] V. Misra, S. Ioannidis, A. Chaintreau, and L. Massoulié, "Incentivizing peer-assisted services: a fluid shapley value approach," *SIGMETRICS Perform. Eval. Rev.*, vol. 38, no. 1, pp. 215–226, Jun. 2010. [Online]. Available: http://doi.acm.org/10.1145/1811099.1811064

[11] Cisco, "Cdn federation," 10 2012, available on http://www.cisco.com/web/about/ac79/docs/sp/CDN-Federation_Phase-2-Pilot.pdf.

[12] W. B. Norton, *The Internet Peering Playbook: Connecting to the Core of the Internet*, 2nd ed. DrPeering Press, December 2012.

[13] N. KAMIYAMA, T. MORI, R. KAWAHARA, and H. HASEGAWA, "Optimally designing isp-operated cdn," *IEICE Transactions on Communications*, vol. E96.B, no. 3, pp. 790–801, March 2013.

[14] IETF, "Cdn interconnect ietf working group," 2013, available on https://datatracker.ietf.org/wg/cdni/. [Online]. Available: https://datatracker.ietf.org/wg/cdni/

[15] "Figaro project," http://www.ict-figaro.eu/, 2012.

[16] M. Cha, P. Rodriguez, S. Moon, and J. Crowcroft, "On next-generation telco-managed p2p tv architectures," in *Proceedings of the 7th international conference on Peer-to-peer systems*, ser. IPTPS'08. Berkeley, CA, USA: USENIX Association, 2008, pp. 5–5. [Online]. Available: http://dl.acm.org/citation.cfm?id=1855641.1855646

[17] W. Jiang, S. Ioannidis, L. Massoulié, and F. Picconi, "Orchestrating massively distributed cdns," in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, ser. CoNEXT '12. New York, NY, USA: ACM, 2012, pp. 133–144. [Online]. Available: http://doi.acm.org/10.1145/2413176.2413193

[18] N. MAKI, T. NISHIO, R. SHINKUMA, T. MORI, N. KAMIYAMA, R. KAWAHARA, and T. TAKAHASHI, "Traffic engineering of peer-assisted content delivery network with content-oriented incentive mechanism," *IEICE Transactions on Information and Systems*, vol. E95.D, no. 12, pp. 2860–2869, December 2012.

[19] B. Mathieu and Y. Levene, "Impact of ftth deployment on live streaming delivery systems," in *Computers and Communications (ISCC), 2012 IEEE Symposium on*, july 2012, pp. 000 259 –000 264.

[20] Y. Borghol, S. Mitra, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti, "Characterizing and modelling popularity of user-generated videos," *Perform. Eval.*, vol. 68, no. 11, pp. 1037–1055, Nov. 2011. [Online]. Available: http://dx.doi.org/10.1016/j.peva.2011.07.008

[21] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.

[22] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Characteristics of youtube network traffic at a campus network - measurements, models, and implications," *Comput. Netw.*, vol. 53, no. 4, pp. 501–514, Mar. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2008.09.022