

A Review of Server Bandwidth Saving Method for Peer-to-Peer Video-on-Demand System

Salehah Hamzah^{1,2*}, Putra Sumari²

¹ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 02600, Arau, Perlis, Malaysia.

² School of Computer Science, Universiti Sains Malaysia, 11800, Minden, Penang, Malaysia.

* Corresponding author. Tel.:+60122322844; email: saleha@perlis.uitm.edu.my

Manuscript submitted July 31, 2015; accepted December 20, 2015

doi: 10.17706/ijcce.2016.5.5.367-373

Abstract: Akamai and Netflix are the most high performance Internet video delivery service provider with their efficient CDN infrastructures. However, due to the high cost of CDN deployment, strong incentives are created to find another lower cost delivery service options. Peer to peer (P2P) network represents less investment architecture that show high potential to reduce the dependency of CDN. A P2P based system shifts the load of content delivery from the content providers' servers to the users. However, high consumption of server bandwidth often become an issue in P2P based streaming as user demand increase rapidly. This paper review two category of methods of server bandwidth saving: (a) peer's bandwidth allocation method and (b) server bandwidth provisioning method which potential for P2P based video on demand (VoD) system. We then conclude our review with CDN-P2P hybrid architectures research towards becoming a high performance video distribution.

Key words: P2P video on demand, peer to peer network, server bandwidth saving, CDN infrastructures.

1. Introduction

The ongoing growth of broadband technology in the worldwide market has been driven by the hunger of customers for various multimedia services. Video, which carries much more information than audio required longer download times and suffers from server overloading when a large number of user demand arrive using traditional client-server model. Large scale video streaming systems such as live streaming and video-on-Demand streaming are two multimedia service which recently attracted a large number of users used to distribute content using well-known Content Delivery Networks (CDN) such as are Akamai¹ and Limelight². Both of the system used to distribute the media contents from the origin server to the replica servers close to the end clients [1]. Another CDN, NetFlix, which has deployed more than 77,000 video servers globally [2] and can stream out High Definition (HD) video with average bitrate reaching to 3.6 Mbps [3]. YouTube³, the most popular on-demand streaming application has more than thousands of users concurrently online to stream for popular videos used 6 large data-centers located in United States to serve videos to users [4].

In fact, CDN benefits professional administrators and fully provisioned resources as they can control and authenticate the content they distribute. However, due to high CDN distribution costs, strong incentives are

¹Akamai, <http://www.akamai.com>.

²LimeLight Networks, <http://www.limelight.com>.

³YouTube, <http://www.youtube.com>.

created to find another lower cost delivery options [5].

Peer-to-peer (P2P) represents less investment architecture that could reduce the dependency of CDN whereby each node in P2P network can be at the same timeserver and client [5]. It means users or peers consuming the video stream also contribute their bandwidth and storage resources to the system by forwarding the video stream to other peers, thus reducing the content provider's cost. P2P network formerly known as file sharing network from the advent Napster music sharing. Researchers have realized their potential of P2P technologies for distribution of large-volume content such as movies and software Commercial services deployed using P2P based system are Skype⁴, BitTorrent⁵ file distribution protocol and also IPTV such as live streaming and on demand streaming. Interested readers can review the details of peer to peer system in [6].

A P2P based system shifts the load of content delivery from the content providers' servers to the users. Thus, it would become the best interest for content providers to reduce their provision cost that is server bandwidth cost. However, high consumption of server bandwidth often become an issue in P2P based streaming which derive from two common problems. First, the dynamic of peer churn that brings insufficient bandwidth supply from peers to serve others in each channel when the demand increases. This is the most common problem of using P2P network. Secondly, a peer must receive the correct sequence of video chunks content before its playback deadline to avoid playback interruptions. The problem is related to the chunks transmission scheduling mechanism which takes longer time searching for peers to locate and download desired chunks content from other peers. Therefore, in this paper, we highlight several potential methods of server bandwidth resource saving over two problems identified above in order to develop an effective P2P Video on-demand (VoD) system.

The rest of the paper is organizes as follows. In Section 2, we briefly discuss overview of P2P Video on-Demand. In Section 3, we present the literature review of saving methods of server bandwidth for P2P VoD system. Finally, in Section 4, we conclude with significant works on CDN-P2P hybrid developments for future video distribution service.

2. Overview of P2P Video-on-Demand

The success of P2P based streaming has shown through the real-world of China-based video streaming such as CoolStreaming⁶, PPLive⁷, UUSee⁸ and PPStream⁹. Basically, the mechanism for P2P based streaming has inspired by the protocol in BitTorrent application developed by Bram Cohen [7] The protocol design mainly used for fast distribution of large files such as movies of software which splitting files into small and fixed size pieces called chunks or blocks to be downloaded by users. Our study focus on Video on-demand system a user driven multimedia streaming service which can perform VCR-like operations (i.e., forward, backward, skip, etc.) and allows users to choose any channels to watch any video at any time [8]. At any moment, VoD users may watching the same video but they are asynchronous which means different users watching different parts of the video. Thus, the playbacks of the same video on different users are not synchronized. A huge size of buffer is needed in order to satisfy the diversified request from peers on different kinds of video programs.

For that reason, bandwidth usage in multichannel Video on Demand (VoD) systems varies rapidly over time, due to unpredictable of user demand that making difficult to keep track of available system resources.

⁴Skype, <http://www.skype.com>.

⁵BitTorrent, <http://www.bittorrent.com>

⁶Coolstreaming, <http://www.coolstreaming.us>.

⁷PPLive, <http://www.pptv.com>.

⁸UUSee, <http://www.uusee.com>.

⁹PPStream, <http://www.ppstream.com>.

Those peers that cannot obtain enough bandwidth from other peers will go to the server as a last resort. The total average server bandwidth utilization is should be reduced. While streaming, if the video stored in the server are always replicated by peer, no additional bandwidth is needed from the servers. However, replication scheme is beyond our scope in this paper as they mainly design to satisfy the playback continuity performance and managing supply-demand relationship. This study highlight double-auction based prefetching, peer's bandwidth prediction and helper-assisted approach as a method to increase peer's upload bandwidth in the system. The second method is server bandwidth provisioning such as demand forecast system, computing maximum flow of server bandwidth and stochastic fluid methodology which aim to manage high demand in large scale P2P VoD system.

3. Review of Server Bandwidth Saving Method

3.1. Peer's Bandwidth Allocation

A very simple solution to minimize server bandwidth consumption dramatically in peer-assisted network is allocated the surplus bandwidth of each peer (i.e., the average upload bandwidth among peers is higher than the streaming rate.) through prefetching strategy [9]. Prefetching in P2P streaming means allocating peer's surplus bandwidth to help peers catch up the missing chunks in the case of peer churn problem. The critical question is how to allocate the instantaneous surplus upload capacity among the peers in the system. To answer the question, water leveling policy is introduced where all users' buffer are treated as water tanks and the buffer with the lowest level is filled first [10].

As peer is asynchronous in real-world peer-assisted VoD systems, optimal prefetching should made practical decision on which segment should be prefetched from which peer. Prefetching in VoD system commonly design to improve the response latency of interactive requests (VCR) such as seek latency (i.e., an interval between the request time and the playback time of segment) [11]-[13]. However, prefetching scheme in [14] shows its potential on server bandwidth resource saving by treating the prefetching behavior as double auction market principle. The approach takes the entire market consists of M sub-markets, with each of them trading one of the M segments. A segment is act as the commodity to be traded in a sub-market. Every segment is associated with a price P_k at each time .Every peer makes their local decisions by participating in these trading markets with their "bids" and "asks" for segments, behaving as both buyers and sellers. A trade is executed immediately when a "bid" is larger than or equal to an "ask". Thus, using double auction principle, prefetching mechanism prove to regulate not only which segments should be uploaded or downloaded, but also to and from whom.

The use of peer bandwidth and storage resources intelligently is the goal of P2P based systems which depends on cooperative strategies among peers in the network. The inter-torrent approach applied in BitTorrent system [15] brought the cooperative design in multichannel P2P streaming where the upload bandwidth from one channel can be used to "help" another channel based the number of seeders within the swarm. Such work [16] trades the seeders surplus bandwidth for downloading chunks in BitTorrent network. Seeders which mostly found in larger sized swarms or popular swarms manage the trading to downloaders in smaller sized swarm which often have insufficient supply bandwidth. However, there exist a special kind of peers who are not interested in downloading file. Thus those peer can act as on-demand cache in a BitTorrent swarm [17]. Several works [18]-[21] attempt to utilize helpers to maximize the streaming capacity for P2P multichannel streaming. The significant work in [22] develop helper-assisted P2P VoD system by optimizes helper's storage and bandwidth with primal dual resource allocation algorithm design and use Markov chain approximation for updating helper's topology according to number of demand. However, the critical task while utilizing helpers is an incentive mechanism which required to encourage peer to join as helper such as reputation based mechanism, game-theoretic model, pricing based

model and taxation mechanism [8].

The common issue in P2P VoD system is peer churn problem that introduces fluctuations over available upload bandwidth in P2P system and cause the server load increase. The solution is bandwidth prediction as an improving of bandwidth allocation to decrease peer churn and further decrease the server load [23]. The allocation prediction takes two scheduling protocol for supplier peer and receiver peer. Supplier peer determine which chunk requests should be served first and then utilize their surplus upload bandwidth to help receiver peers to prefetch chunks. A receiver peer must prioritize its requests for its missing chunks while supplier peer which contribute their upload bandwidth differentiates receiver peer according to their stability and playback position. A new predictor is designed to identify the potential stable peer. It predicts the stability of a peer for remaining online based on the propose ranking procedure. A predicted peer depends on the predefined threshold. A peer will leave if it has been in the session for fewer than the predefined threshold. The performance of the prediction method manage to improve about 8% of decreasing server load compared to traditional bandwidth allocation method and centralized bandwidth allocation method.

3.2. Server Bandwidth Provisioning

The bandwidth prediction scheme in [23] aims to increase surplus bandwidth of peer. However, capacity planning or prediction system over server bandwidth would be a best choice towards for satisfying time varying video popularity for large scale streaming system. Server capacities have increasingly become a bottleneck in real-world P2P live streaming solutions [24]. For that reason, the authors propose server bandwidth provisioning scheme named Ration has developed for P2P live streaming using time series forecasting and dynamic regression technique. Based on the predicted channel popularity and the streaming quality function for each channel, the amount of server bandwidth can be adjust and share to each channel can be allocated by enhanced water-leveling approach that enhance the computation of capacity without from the beginning. The approach can guarantee that the server capacity allocated to each channel may not exceed the amount $q_c = 1$. The value of q_c denote the streaming quality of channel c , (i.e., the percentage of high-quality peers in the channel that have a buffer count of more than 80% of the size of its playback buffer). Ration has been implemented in streaming servers in a multi-ISP mesh-based system which replays real-world streaming scenarios. The result shows Ration is potentially could provide guideline for deployment of server capacities in multiple channel for each ISP.

Previous work [25] used stochastic fluid approach to achieve condition of universal streaming in P2P live streaming. In contrast, the work in [26] use stochastic fluid methodology to derive analytical approximations of the bandwidth requested from servers for developing self-sustainability condition in P2P video on-demand. This approach accounts for several important effects such as peer bandwidth heterogeneity, peer churn, non-sequential chunk transmission. Gaussian approximation is used to characterize the server bandwidth distribution from sequential chunks delivery and non-sequential chunks delivery. The result shown under stationary conditions, server bandwidth performance achieves a maximum as the number of users increase, and then decreases to zero provided that $T > T_d$ (i.e., the average activity time is larger than the average download time).

In non-stationary condition, the system performance depends on the relationship between the average peer upload bandwidth U and the download rate when $U < d_v$. If value $U < d_v$, the bandwidth demand from server scales linearly with the system size. If $U > d_v$ the bandwidth demanded from the servers is essentially independent from number of watching users.

The third method in our review is automated demand forecasting scheme which developed to predict the aggregate bandwidth demand in each video channel [27]. The system using time-series modeling technique named seasonal ARIMA (autoregressive integrated moving-average) focuses on predicting non-stationary

opulation series that is during flashcrowd. The inference of initial demand evolution pattern which has retrieved from several flashcrowd channels is model by mixture of Gaussians. The results show the scheme able to forecast about 2.5 hours ahead of time. Finally, the method which takes pushing the maximum reduction of the server bandwidth consumption [28] which derive from computing the maximum capacity of available resources using maximum flow problem approach which further discussed in [29].

4. CDN-P2P Hybrid Architecture

Due to the maturity of P2P system and high performance CDN system to support large scale streaming, research communities have come to the challenge of combining the P2P architecture and CDN system. Early CDN-P2P-hybrid architectures has been presented in [30]-[32] which show the hybrid architecture is highly cost-effective .Related works of hybrid architecture model of CDN-P2P systems and the integration issues of CDN and P2P has been discussed in [33]. A solid starting point for developing a hybrid CDN architecture by extending the topology CDN with P2P overlay networks into three different domains named CDN provider domain, CDN customer domain and end user domain [34].Other hybrid architecture [35] designed support P2PVoD system by developing ISP-friendly overlay network and PeerCDN as potential architecture in Metropolitan Area Network (MAN)[36]. Livesky is one of the commercial hybrid CDN-P2P architecture [37] which successfully saves more than 40% of the CDN bandwidth cost by integrating P2P technology. Furthermore, Livesky shows more than 85% of the clients wait less than 15s for playback to commence from clicking on the hyperlink which is significantly better than traditional P2P streaming systems. A significant study in [38] shows hybrid P2P-CDN system able to cut the server load for both Akamai and Limelight by more than 2/3 in Internet video-on-demand; an even if peers only upload at limited rate (1/3 of their upload bandwidths). To conclude about CDN P2P hybrid system, the authors [39] has studied the architecture NetSession system, a large commercial hybrid CDN P2P called peer-assisted CDN that operated by Akamai provider. The study delivers the keys of potential benefits in CDN P2P hybrid architecture such as peer-assistance for improving CDN global coverage and the few risks such as traffic impact of ISPs.

References

- [1] Peng, G. (2004). CDN: Content distribution network. Retrieved May 30, 2014, from <http://arxiv.org/abs/cs.NI/0411069>
- [2] Woolley, S. (November 2010). Netflix will ruin the internet! Retrieved May 28 2014, from <http://tech.fortune.cnn.com/2010/11/04/netflix-will-ruin-the-internet/>
- [3] Adhikari, V. K., Guo, Y., Hao, F., Varvello, M., Hilt, V., Steiner, M., & Zhang, Z. (2012). Unreeling Netflix: Understanding and improving multi-CDN movie delivery. *Proceedings of 2012 IEEE INFOCOM*.
- [4] Adhikari, V. K., Jain, S., & Zhang, Z. (2011). Where do you "Tube"? Uncovering YouTube server selection strategy. *Proceedings of 20th International Conference on Computer Communication and Network*.
- [5] Wildman, I. H. S. S., & Bauer, J. M. (2010). P2P, CDNs and hybrid networks: The economics of internet video distribution. *International Telecommunications Policy Review*, 17(4), 1-22.
- [6] Brands, E. H. T. B., & Karagiannis, G. (2009). Taxonomy of P2P applications. *Proceedings of IEEE Globecom Workshop*. Honolulu.
- [7] Cohen, B. (2003). Incentives build robustness in BitTorrent. *Proceedings of Workshop on Economics of Peer to Peer Systems*.
- [8] Ramzan, N., Park, H., & Izquierdo, E. (2012). Video streaming over P2P network: Challenges and opportunities. *Signal Processing: Image Communication*, 27(5), 401-411.
- [9] Huang, Y., Fu, T. Z. J., Chiu, D-M., Lui, J., & Huang, C. (2008). Challenges, design and analysis of a large-scale P2P-VOD system. *ACM SIGCOMM Computer Communication Review*, 38, 375–388.

- [10] Huang, C., Li, J., & Ross, K. W. (2007). Can internet video-on-demand be profitable? *ACM SIGCOMM Computer Communication Review*, 37, 133–144.
- [11] Xu, T., Wang, W., Ye, B., Li, W., Lu, S., & Gao, Y. (2009). Prediction-based prefetching to support VCR-like operations in gossip-based P2P VoD systems. *Proceedings of 15th International Conference on Parallel and Distributed Systems*. Shenzhen, China.
- [12] Abboud, O., Pussep, K., Mueller, M., Kovacevic, A., & Steinmetz, R. (2010). advanced prefetching and upload strategies for P2P video-on-demand. *Proceedings of ACM Workshop on Advanced video Streaming techniques for Peer-to-Peer Networks and Social Networking*. Firenze, Italy.
- [13] He, Y., & Guan, L. (2009). Prefetching optimization in P2P VoD applications. *Proceedings of First International Conference on Advances in Multimedia*. Colmar, France.
- [14] Feng, Y., Li, B., & Li, B. (2010). Peer-assisted VoD prefetching in double auction markets. *Proceedings of 18th IEEE International Conference on Network Protocols* (pp. 275-284). Kyoto, Japan.
- [15] Guo, L., Chen, S., Xiao, Z., Tan, E., Ding, X., & Zhang, X. (2007) A performance study of BitTorrent-like peer-to-peer systems. *IEEE Journal Selected Areas in Communication*, 25(1), 155-169.
- [16] Lee, H., Yoshida, M., & Nakao, A. (2011). SeederTrading: Trading swarm capacity for improving content distribution. *Proceedings of IEEE International Symposium of Parallel and Distributed Processing Workshop and PhD Forum* (pp. 1587-1594). Shanghai, China.
- [17] Wong, J. H. T. (2004). *Enhancing Collaborative Content Delivery with Helpers*. Master's Thesis, University of British Columbia.
- [18] He, Y., & Guan, L. (2010). Solving streaming capacity problems in P2P VoD systems. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11), 1638–1642.
- [19] Wang, Z., Wu, C., Sun, L., & Yang, S. (2010). Strategies of collaboration in multi-channel P2P VoD streaming. *Proceedings of IEEE Global Telecommunication Conference*. Miami, Florida, USA.
- [20] Mostafavi, S. A., & Dehghan, M. (2014). Decentralized adaptive helper selection in multi-channel P2P streaming systems. *Proceedings of IEEE International Conference on Distributed Computing System Workshops*. Madrid, Spain.
- [21] Zhang, H., et al. (2009). Scaling peer-to-peer video-on-demand systems using helpers. *Proceedings of 16th IEEE International Conference on Image Processing* (pp. 3053-3056), Cairo, Egypt.
- [22] Zhang, H., Chen, M., Parekh, A., & Ramchandran, K. (2010). An adaptive multi-channel P2P video-on-demand system using plug-and-play helpers. Retrieved August 23 2014, from <http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-111.pdf>
- [23] Chen, Z., Xue, K., Hong, P., & Lu, H. (2009). Differentiated bandwidth allocation for reducing server load in P2P VOD. *Proceedings of Eight International Conference on Grid and Cooperative Computing* (pp. 31-36). Lanzhou, Gansu, China.
- [24] Wu, C., Li, B., & Zhao, S. (2008). Multi-channel Live P2P streaming: Refocusing on servers. *Proceedings of 27th IEEE Conference of Computer Communications*. Phoenix, Arizona, USA.
- [25] Kumar, R., Liu, Y., & Ross, K. (2007). Stochastic fluid theory for P2P streaming systems. *Proceedings of 26th IEEE International Conference on Computer Communications* (pp. 919-927). Alaska, USA.
- [26] Ciullo, D., Martina, V., Garetto, M., Leonardi, E., & Torrisi, G. L. (2012). Stochastic analysis of self-sustainability in peer-assisted VoD systems. *Proceedings of IEEE International Conference on Computer Communications* (pp. 1539-1547). Orlando, Florida.
- [27] Niu, D., Liu, Z., Li, B., & Zhao, S. (2011). Demand forecast and performance prediction in peer-assisted on-demand streaming systems. *Proceedings of 2011 IEEE INFOCOM* (pp. 421-425). Shanghai, China.
- [28] Xu, K., Wang, H., Liu, J., Lin, S., & Xu, L. (2014). Pushing server bandwidth consumption to the limit: Modeling and analysis of peer-assisted VoD. *IEEE Transaction on Network and Service Management*,

- 11(4), 472-485.
- [29] Goldberg, A., & Tarjan, R. (1988). A new approach to the maximum flow problem. *Journal of ACM*, 35(4), 921-940.
- [30] Xu, D., Chai, H.-K., Rosenberg, C., & Kulkarni, S. (2003). Analysis of a hybrid architecture for cost-effective streaming media distribution. *Proceedings of SPIE Conference on Multimedia Computing and Networking*.
- [31] Xu, D., Kulkarni, S., Rosenberg, C., & Chai, H-K. (2006). Analysis of a CDN-P2P hybrid architecture for cost-effective streaming media distribution. *Multimedia Systems*, 11(4), 383-399.
- [32] Hafeeda, M., Bhargava, B., & Yau, D. (2004). A hybrid architecture for cost-effective on-demand media streaming. *Journal of Computer Networks*, 44(3), 353–382.
- [33] Lu, Z., Wang, Y., Yang, Y., & Yang, Y. R. (2012). An analysis and comparison of CDN-P2P hybrid content delivery system and model. *Journal of Communications*, 7(3), 232-245.
- [34] Pakkala, D., & Latvakoski, J. (2005). Towards a peer-to-peer extended content delivery network. Retrieved 20 June 2015, from <http://www.eurasip.org/Proceedings/Ext/IST05/papers/99.pdf>
- [35] Lu, Z., Wu, J., Chen, L., Huang, S., & Huang, Y. (2010). CPH-VOD: A novel CDN-P2P hybrid architecture based VoD scheme. In L. Chen, P. Triantafillou, & T. Suel (Eds.), *Web Information System Engineering* (pp. 578-586). Springer Berlin Heidelberg.
- [36] Wu, J., Lu, Z., Liu, B., & Zhang, S. (2008). PeerCDN: A novel P2P network assisted streaming content delivery network scheme. *Proceedings of 8th IEEE International Conference on Computer and Information Technology*. Sydney, Australia.
- [37] Yin, H., Liu, X., Zhan, T., Sekar, V., Qiu, F., Lin, C., Zhang, H., & Li, B. (2009). Design and deployment of a hybrid CDN-P2P system for live video streaming: Experiences with LiveSky. *Proceedings of 17th ACM international Conference on Multimedia*. Beijing, China.
- [38] Huang, C., Wang, A., Li, J., & Ross, K. W. (2008). Understanding hybrid CDN-P2P: Why limelight needs its own red swoosh. *Proceedings of 8th International Workshop on Network and Operating Systems Support for Digital Audio and Video*. Braunschweig, Germany.
- [39] Zhao, M., et al. (2013). Peer-assisted content distribution in Akamai Netsession. *Proceedings of 2013 Conference on Internet Measurement*. Barcelona, Spain.



Salehah Hamzah was born in Perak, Malaysia in 1975. She received her bachelor degree in computer science from Universiti Teknologi Malaysia, Johor, Malaysia in 1998 and the master degree in information technology from Universiti Utara Malaysia, in 2005. Currently, she is a PhD candidate in School of Computer Science, Universiti Sains Malaysia, Penang, Malaysia under the supervision of Assoc. Prof. Putra Sumari from School of Computer Science USM. Her research interest is in the areas of multimedia networking and peer to peer computing.



Putra Sumari received his masters and PhD degrees in 1997 and 2000 from Liverpool University, England, respectively. Currently, he is an associate professor at the School of Computer Science, Universiti Sains Malaysia, Penang, Malaysia. He is also a member of ACM and IEEE, program committee and reviewer of several journals and International Conferences. He has published more than hundred papers including journal and conferences. His research areas are multimedia communication and networking, specifically on video-on-demand system and content distribution network and content based information retrieval.