Challenges on Peer-to-Peer Live Media Streaming

Wen Gao and Longshe Huo

Institute of Digital Media,
Peking University, China
{wgao, lshuo}@pku.edu.cn
http://idm.pku.edu.cn

Abstract. Recently there has been significant interest in the use of peer-to-peer technologies for deploying large-scale live media streaming systems over the Internet. In this position paper, we first give a brief survey on the state-of-the-art of peer-to-peer streaming technologies, and then summarize and discuss some major challenges and opportunities in this research area for future directions.

Keywords: Peer-to-Peer, media streaming.

1 Introduction

During the past decade, the rapid development of the Internet has changed the conventional ways that people access and consume information. Besides traditional data services, various new multimedia contents would also be delivered over the same IP network, among which the live media streaming service will play a more important role. The typical applications of live media streaming include Internet television, distance education, sports events broadcasting, online games, and etc.

Recently, Peer-to-Peer (P2P) has emerged as a promising technique for deploying large-scale live media streaming systems over the Internet, which represents the paradigm shift from conventional networking applications. In a P2P system, peers communicate directly with each other for the sharing and exchanging of data as well as other resources such as storage and CPU capacity, each peer acts both as a client who consumes resources from other peers, and also as a server who provides service for others. Compared with traditional streaming techniques such IP multicast and CDN (content delivery networks), P2P based streaming system has the advantages of requiring no dedicated infrastructure and being able to self-scale as the resources of the network increase with the number of users.

A great number of systems for P2P based live media streaming have been proposed and developed in recent years. Merely in China, nowadays there are about more than a dozen of P2P streaming applications deployed in the Internet. In this position paper, we first give a brief survey on the state-of-the-art of P2P streaming technologies, and then summarize and indicate some major challenges and opportunities in this research area for future directions.

N. Sebe, Y. Liu, and Y. Zhuang (Eds.): MCAM 2007, LNCS 4577, pp. 37 – 41, 2007.

2 State-of-the-Art of P2P Streaming Technologies

In typical P2P streaming implementations, media data are distributed along a application-layer logical overlay network constructed over the underlying physical IP network. To construct and maintain an efficient overlay network, mainly three questions should be answered. The first relates to the P2P network architecture, i.e., what topologies should the overlay network be constructed? The second concerns routing and scheduling of media data, i.e., once the overlay topology is determined, how to find and select appropriate upstream peers from which the current peer receives the needed media data? The third is membership management, i.e., how to manage and adapt the unpredictable behaviors of peer joining and departure?

Recently, several P2P streaming systems and algorithms have been proposed to address the above issues. From the view of network topology, current systems can be classified into three categories approximately: tree-based topology, forest-based topology, and mesh topology. Following we outline a brief survey on P2P streaming techniques according to this classification.

(1) Tree-based topology
PeerCast [1] is a typical example of tree-based P2P streaming systems. In PeerCast, peers are organized into multicast trees for delivering data, with each data packet being disseminated using the same structure. When a peer receives a data packet, it also forwards copies of the packet to each of its children. Since all data packets follow this structure, it becomes critical to ensure the structure is optimized to offer good performance to all receivers.

Generally, there exist four route selection strategies in tree-based topology: random selection, round-robin selection, smart selection according to physical placement, and smart selection according to bandwidth. To achieve a balanced multicast tree, custom routing policy should be chosen carefully for individual peer node.

(2) Forest-based topology
Conventional tree-based multicast is inherently not well matched to a cooperative environment. The reason is that in any multicast tree, the burden of duplicating and forwarding multicast traffic is carried by the small subset of the peers that are interior nodes in the tree. Most of the peers are leaf nodes and contribute no resources. This conflicts with the expectation that all peers should share the forwarding load.

To address this problem, forest-based architecture is beneficial, which constructs a forest of multicast trees that distributes the forwarding load subject to the bandwidth constraints of the participating nodes in a decentralized, scalable, efficient and self-organizing manner. A typical model of forest-based P2P streaming system is SplitStream [2]. The key idea of SplitStream is to split the original media data into several stripes, and multicast each stripe using a separate tree. Peers join as many trees as there are stripes they wish to receive, and they specify an upper bound on the number of stripes that they are willing to forward. The challenge is to construct this forest of multicast trees such that an interior node in one tree is a leaf node in all the remaining trees and the bandwidth constraints specified by the nodes are satisfied. This ensures that the forwarding load can be spread across all participating peers.

(3) Mesh topology

In conventional tree-based P2P streaming architectures, at the same time a peer can only receive data from a single upstream sender. Due to the dynamics and heterogeneity of network bandwidths, a single peer sender may not be able to contribute full streaming bandwidth to a peer receiver. This may cause serious performance problems for media decoding and rendering, since the received media frames in some end users may be incomplete.

In forest-based systems, each peer can join many different multicast trees, and receive data from different upstream senders. However, for a given stripe of a media stream, a peer can only receive the data of this stripe from a single sender, thus results in the same problem like the case of single tree.

Multi-sender scheme is more efficient to overcome these problems. In this scheme, at the same time a peer can select and receive data from a different set of senders, each contributing a portion of the streaming bandwidth. In addition, members of the sender set may change dynamically due to their unpredictable online/offline statuses. Since the data flow has not a fixed pattern, every peer can send and also receive data from each other, thus the topology of data plane likes mesh. The main challenges of mesh topology are how to select the proper set of senders and how to cooperate and schedule the data sending of different senders.

Recently, DONet [3] implemented a multi-sender model by introducing a simple and straightforward data-driven design, which does not maintain an even more complex structure. The core of DONet is the data-centric design of streaming overlay, and the Gossip-based message distribution algorithm. In DONet, a node always forwards data to others that are expecting the data. In other words, it is the availability of data but not a specific overlay structure that guides the flow directions. Such a design is suitable for overlay with high dynamic nodes. Experiments showed that, compared with a tree-based overlay, DONet can achieve much more continuous streaming with comparable delay.

3 Technical Challenges and Opportunities

Though some successes have been made in recent years, especially with the introducing of mesh-based approaches, there are still challenging problems and open issues need to be overcome in P2P live media streaming.

The main problem results from the heterogeneity of the underlying IP networks. There exist mainly two types of heterogeneities in the current Internet: heterogeneous receivers and asymmetric access bandwidths. In a P2P based live media streaming system, for each individual peer the receiving capability is decided by its downlink bandwidth, however for the whole system the total available bandwidth is decided by the sum of the uplink bandwidths of all the participated peers. Under this situation, same and perfect QoS is hard to be guaranteed for all of the participated peers. For example, if the access bandwidth of a peer is less than the average bit rate of the media stream it requires, or the sum of the uplink bandwidths of all upstream peers who provide data for this peer is less than the average bit rate, then random packet-losses may occur ether during the network or at the buffer of upstream peers. This may lead to incorrect decoding at the client side even partial data have been received, which means not only the waste of bandwidth resources, but also degraded media reconstruction qualities.

The most hopeful solution to this problem is to provide self-adaptive QoS for each individual peer according to the current network conditions, at the same time the total available uplink bandwidths of all peers are utilized as full as possible. To satisfy this objective, three main issues should be addressed:

(1) Content aware media data organization

Current P2P streaming systems focus mainly on network topology and protocol design, but pay rare attention to the media contents carried over the network. In fact, since streaming media have their distinct characteristics from normal data file, good performance can be achieved only when both the characteristics of media coding and networking are considered together perfectly. While scalable coding techniques hold promise for providing network adaptive media transmission, they are yet to be deployed in today's mainstream media codec. A promising solution is to partition the current non-scalable coded media data based on content analysis, and reorganize them into another form with scalable capability to some extent, so that selective and priority-based schedule strategies can be used while transmission.

(2) Priority-based media data delivery mechanism

For the quasi-scalable media data prepared above, efficient transmission and control mechanisms should be invented to guarantee that the minimal decodable media units (for example, a video frame or slice) can be transmitted to the receiver in a restrict order based on their priorities. This implies that every data unit received by a peer at any time is intact and decodable independent of any still un-received media data. By this way, no waste of bandwidth is involved and free-error and fluent media experience can be obtained even in the case of worse network conditions.

(3) QoS adaptive multi-source and layered media data schedule algorithm

Based on the above content aware data organization and priority-based delivery mechanism, efficient data schedule algorithms are needed to retrieve data from multiple senders in order to maximize the overall bandwidth utilization of the whole network and minimize the average media reconstruction distortion of all users. Compared with conventional P2P streaming systems which simply partition a streaming media into a series of data blocks and schedule each data block as the minimal transmitting unit, the scheduling model of this system and its solutions are more complicate to establish and resolve.

In another paper of this workshop [4], we present a novel data organizing and delivery framework for P2P live media streaming, which takes into account both the characteristics of media coding and P2P networking, and aims to provide self-adaptive QoS for different users of the Internet. We hope that it can become a good start for inspiring more and more researches on beneficial solution to the above problems.

4 Conclusion

Recently, P2P streaming has attracted a lot of attentions from both academy and industry. Various P2P media streaming algorithms have been studied, and the systems have been developed. However, a number of key technical challenges still need to be overcome in order to maximize the whole network resource utilizations. Opportunities are always along with challenges. We still have a long way to go in the research area of P2P based live media streaming systems.

References

1. Deshpande, H., Bawa, M., Garcia-Molina, H.: Streaming live media over a Peer-to-Peer network, Stanford database group technical report (2001-20) (August 2001)
2. Castro, M., Druschel, P., Kermarrec, A.-M., Nandi, A., Rowstron, A., Singh, A.: Split-Stream: High-bandwidth content distribution in a cooperative environment, In: Proc. the International Workshop on Peer-to-Peer Systems, Berkeley, CA (February 2003)
3. Zhang, X., Liu, J., Li, B., Yum, T.: Coolstreaming/DONet: A data-driven overlay network for live media streaming. Proc. IEEE INFOCOM ’05 (2005)
4. Huo, L., Gao, W., Fu, Q., Guo, R., Chen, S.: QoS adaptive data organizing and delivery framework for p2p media streaming. In: Proc. MCAM’07 (June 2007)