

A Systematic Study of P2P Streaming Media Based on Cloud Computing

Xiaolai ZHOU

Information Centre
 Beijing Jiaotong University
 Beijing, China
 e-mail: xlzhou@bjtu.edu.cn

Abstract—This study expounds on the integration of P2P technology and cloud computing technology in cloud environment and put forwards the P2P streaming media system resource-sharing network structure based on cloud computing, aiming to provide a new approach for transmitting online video in the age of big data. Potential integration scheme of P2P technology and cloud computing technology is demonstrated in this thesis to find out the optimum approach for the perfect integration.

Keywords-P2P; cloud computing; streaming media; node; cluster; Qos

I. INTRODUCTION

With the rapid development of the internet, broadband application in the future will be dominated by streaming media applications such as distance teaching, video on demand, video conference and telemedicine, providing information and entertainment combined with text, video and audio to users. Streaming media network system must possess high-broadband, high real-time performance and high reliability to satisfy the large-scale internet user groups who want to obtain high-quality streaming media services. In this regard, P2P streaming media transmission based on cloud computing offers a new idea for streaming media network transmission in the age of big data.

II. THE NECESSITY OF INTEGRATION OF CLOUD COMPUTING AND P2P NETWORK

A. P2P Network

Featured by advantages such as high scalability, easy-to-arrange and low consuming costs, P2P streaming media has become the most popular scheme in large-scale network streaming media application implementation in recent years. P2P streaming media consists of two categories i.e. live media and on-demand media. Many P2P streaming media systems have already been extensively applied, among which PPLive, PPStream and SopCast are typical examples. In streaming media system, video files occupy a large proportion of broadband during transmission due to their large data size. If client-server (C/S) mode is adopted, as shown in Figure 1, the total capacity and resources of network are concentrated on the central server, and it will be difficult for centralized computing and storage framework to integrate data between different servers in the client. Under

the circumstance of high concurrency requirements, servers can be the bottleneck of the system due to limited computing capacity, storage, broadband and other resources.

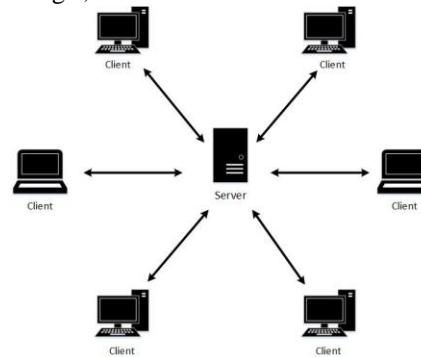


Figure 1. C/S mode

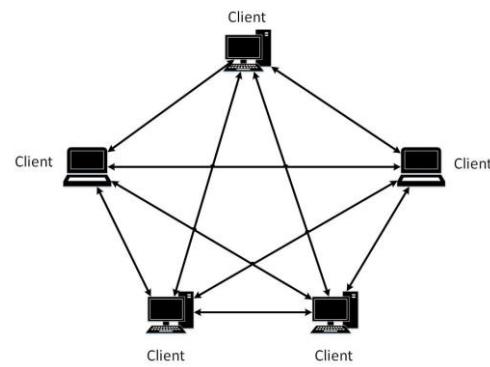


Figure 2. P2P mode

Regarding to the existing problems in C/S mode, P2P technology needs to be adopted to improve scalability, service quality and other performances of streaming media system, as shown in Figure 2. According to P2P technology, all the nodes in the system are equal, both as users and providers of services. By applying P2P technology in streaming media transmission, every node in the system can receive and transmit streaming media data, thus improving data concurrency and overcoming malfunction caused by single point failure in the traditional C/S mode. Main applications of P2P technology include instant messaging software, file sharing system, distributed computing, collaboration work, information retrieval among others. However, the service provided by P2P is unreliable, since

nodes in the overlay network may join or exit at any time, the uploading bandwidth of each node is unpredictable and application programs consume much network broadband, easily causing network congestion.

B. Overview of Cloud Computing

Cloud computing is a mode, which puts ubiquitous, convenient and demand-assigned network access in a sharing and configurable computing resource pool, such as network, servers, storage and application programs and services. These computing resources can be allocated and released quickly, which needs only a small amount of management and interaction with the server, as shown in Figure 3. In cloud computing platform, the data, services, application software, development platform and other resources used by users are all from the cloud computing service provider on the internet. The service provider is responsible for allocating all kinds of resources distributed on the network, load balancing, software arrangement and security control. The user can apply for certain quota of resources according to his own demand, and can adjust the quota according to the change of his demand. In this way, the user no longer needs to buy a large amount of resources for instant peak; instead, he can increase the applied quota prior to the peak, dynamically adapting to the changing demand. Cloud computing makes the acquisition and management of resources convenient for users and reduce their cost demand.

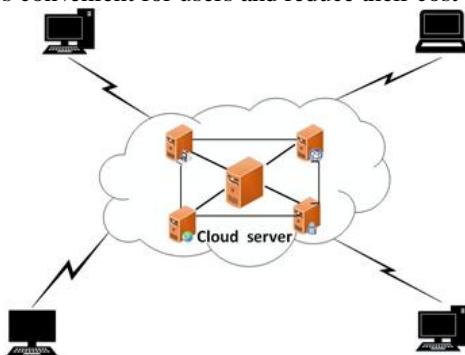


Figure 3. Structural diagram of cloud computing

Elastic scalability and adaptivity of cloud computing is a great advantage for streaming media, such as CALMS (Cloud-Assisted Live Media Streaming). When users worldwide watch the same channel, the peak time varies in different time zones. In such case, CALMS adapts to the demand of streaming media users in time and space by adaptively renting and subtly adjusting cloud server resources, and put forwards the optimal solution to settle the renting cost for cloud servers of different sizes. It is a two-layer structure, namely cloud layer and user layer. The cloud layer consists of streaming media live sources and dynamic renting cloud servers. When a user sends out a request, the cloud layer will redirect a suitable cloud server to him in accordance with regional differences and time demands. Meanwhile, the cloud layer will also adjust the number of rented servers and regional distribution according to demands.

Compared with traditional services based on local devices, cloud service is featured by large scale, high reliability, universality, high scalability, on-demand service and low cost. For common users, they can get services through cloud platform easily and fast at low cost. Basic P2P streaming media transmission system can be effectively optimized by utilizing cloud services, which are accessible in areas covered by network.

C. Analysis of Integration of Cloud Computing and P2P Streaming Media Framework

P2P is suitable to be the main form of streaming media transmission in terms of streaming media transmission framework, despite some disadvantages, such as unpredictability of upload bandwidth of each node and no guarantee on quality of services (QoS) for users. Advantages of cloud computing including reliability of services, high scalability, guarantee on quality of services and charge-on-demand, offer opportunities for cloud-based low-delay and large-scale P2P streaming media transmission.

First, cloud computing environment can provide computing, storage and reliable broadband resources needed by large-scale P2P streaming media live servers. Second, cloud computing technology possesses of dynamic scalability and environment awareness and can be used to make dynamic adjustment of system resources and data transmission based on operation situation monitoring. It is suitable for dynamic P2P peer-to-peer network environment, and can change the research status of unknowingness and uncontrollability of P2P streaming media server resources.

P2P technology can be applied in cloud computing system to design streaming media platform based on cloud computing, make the most of the high reliability of cloud computing to offset the disadvantage of resources unreliability, and resolve resource-related bottleneck problems in the same cloud system where service provider accesses broadband. Meanwhile, waste of idle infrastructure can be avoided thanks to cloud computing scalability, thus realizing efficient utilization of resources, and time consumption and cost caused by geographic discrepancy of the service provider's different data centers can also be reduced. Therefore, the controllability and elastic resources can well remedy deficiencies of P2P network.

III. CLOUD COMPUTING AND P2P STREAMING MEDIA INTEGRATION MODES

Cloud service and P2P technology can be integrated in the following two modes:

A. P2P-in-Cloud Mode

By joining nodes to an existing cloud computing platform, P2P technology can be used to improve the scalability of the cloud computing platform, thus reducing costs of cloud computing resources, storage resources and broadband, by sharing workload between nodes, while maintaining availability of cloud services. P2P-in-Cloud Mode is mainly used in fields such as video-on-demand system, content-based management and file distribution system.

- P2P technology can be used to optimize topological structure of cloud to eliminate cloud system bottlenecks. For instance, there is a centralized master server in the GFS of Google, in which metadata of all file systems are stored, providing index for all the chunk servers and controlling the operation of the whole system. The master server is highly likely to be a bottleneck of the whole system, affecting system scalability. P2P technology can be used to eliminate such a bottleneck. The solution is as follows: directly joining all the chunk servers by means of DHT technology to replace the master server in the GFS, or building DHT network by using a group of master servers to replace single master server in the GFS.
- P2P technology can be used to build the epitaxial structure of cloud, participate in cloud building and maintenance by means of user nodes, and share functions of cloud to reduce costs of the service provider. However, frequent dynamic changes of the participating nodes will result in instability of topological structure in the epitaxial network, thus affecting robustness of the route and the availability of the content. There remains problem of consistency of duplicates since multiple duplicates of cloud data exist in epitaxial user nodes. Moreover, problems such as load balancing and privacy protection also exist due to epitaxial nodes sharing functions of cloud.

B. Cloud-in-P2P Mode

By introducing cloud computing technology into the existing P2P network, reliable resources can be provided to nodes through cloud computing, ensuring high reliability of services in case of node failure. Relatively stable and reliable topological structure can be built with cloud as the super node in P2P network. The super node offered by cloud computing platform can provide relatively stable superstructure and services despite frequent dynamic changes of user nodes. Besides, the super node offered by cloud computing platform can serve as a reliable third party to improve the creditworthiness and security of the whole P2P network. For instance, cloud can be the edge register in P2P streaming media services to improve the speed of multimedia resources distribution and reliability of services. Cloud-in-P2P Mode is mainly applied in fields such as batch synchronization content distribution system and online back system.

IV. P2P STREAMING MEDIA C-P2P TRANSMISSION FRAMEWORK BASED ON CLOUD COMPUTING

With the constant emerging of mass streaming media data on the internet, cloud services are needed to optimize the traditional structural, nonstructural and mixed structural P2P video streaming transmission frameworks. The C-P2P controllable mixed streaming media transmission mode combined P2P streaming media transmission mode with cloud service, is a dynamic combination of nodes. Communication between these nodes are realized by

constantly exchanging messages, which can not only meet users' QoS demand to obtain low-delay and high-quality video, but also help the video provider save costs and the cloud service provider make the best of resources.

A. Framework Description

1) Nodes

- User nodes which are conducting peer-to-peer services, such as terminal computer which is playing video.
- Active nodes and passive nodes, i.e. some computing and storage computers rented from cloud services.
Active node (AN) is a automatic virtual machine composed of single or multiple core variable memory and permanent memory space. Its task is to receive video streaming data from the origin server and immediately distribute these data to user nodes with all its uploading capacity, to ensure user nodes obtain cached video data as early as possible.
Passive node (PN) rents simple cloud storage service. Media source can push any data to PN (known as PUT) and user nodes can fetch data from PN (known as GET). The role is to request PN to push data directly to user nodes in case user nodes fail to obtain video data prior to the qualification time.
- Media origin server is responsible for distributing the starting point of video streaming.

2) Cluster

Cluster refers to the assemble of all user nodes. One node only communicate with some other nodes (known as neighbor), which are subsets of cluster nodes. Cluster undergoes dynamic changes at all times, since new user nodes may join and user nodes in the cluster may voluntarily exit or lose connection at any time.

3) Data block

It is supposed that streaming media source produces video streaming with constant code rate and then divide them into many small blocks of data. Each data block is to have correct play order within the actual time during which video streaming is produced, since not all the data block arrive in the play order.

B. Framework Model

• Benchmark model

Benchmark model can be simply regarded as P2P streaming media transmission framework with a video origin server. PN is introduced on the basis of ordinary P2P video streaming transmission framework, as shown in Figure 4. In the push-pull model, nodes keep exchanging their data validity and use such information to program if they pull needed data block from the neighbor. Nodes have priority to access the data block closest to play order.

Besides sending data block to cluster, video media source is also responsible for storing data in PN. Root node is the first batch of user nodes to receive video data block from the video source in the cluster. Black node refers to user nodes that can receive data block prior to buffer time. Hollow node

is user nodes that obtain data by requesting PN because it is unable to receive data timely and QoS cannot be guaranteed.

As one of the earliest resolutions to guarantee QoS in traditional P2P video streaming transmission framework, benchmark model is too costly. Compared with obtaining data from cluster peer-to-peer nodes, an excess of data block is obtained from PN, which immensely increases storage cost of cloud services and rents too much broadband. Nevertheless, this is not only bandwidth problem. Video streaming timeliness and limits of QoS cannot wait for each node to be optimized in accordance with broadband optimization transmission network, nor enable each node transmit, receive data and information successfully, let alone the network framework is undergoing real-time dynamic changes.

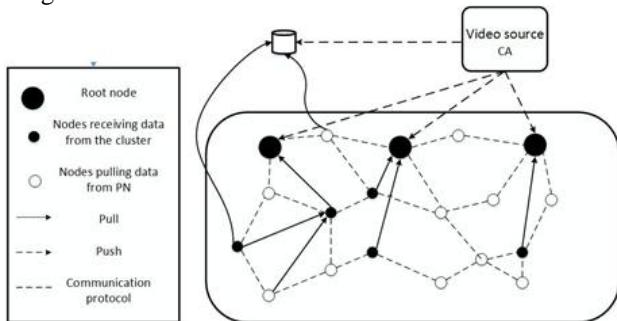


Figure 4. Benchmark model

- Enhancement Model

A certain amount of AN are introduced to the cluster in Enhancement Model, as shown in Figure 5, to ensure that more nodes can receive data block timely. AN can obtain data through media source or other AN, then push the data to other AN or nodes in the cluster. Root nodes of AN need no change unless they lose efficacy, exit from the system or AN find nodes with stronger uploading capacity. To prevent root nodes from receiving repeated video data block, one node can serve as the root node of only one AN, while PN continue providing on-demand video data block. The usage rate of PN declines greatly compared with Benchmark Model. In terms of structure, efficiency of the whole network will be improved in that more nodes receive video streaming data earlier.

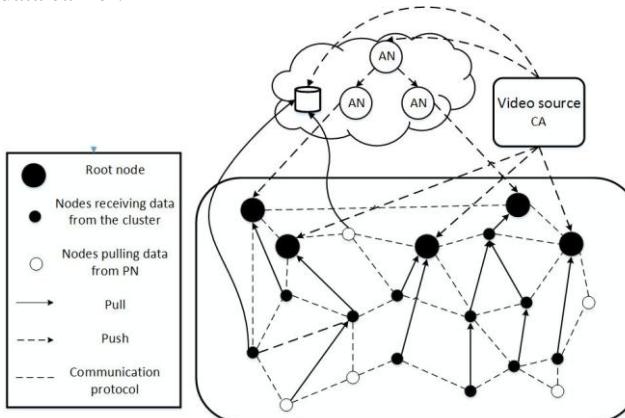


Figure 5. Enhancement model

There are two feasible modes regarding to how to organize AN and push data to them.

- Flat mode: all AN receive data directly from the media source and then transmit these data to each node. The advantage of flat mode lies in shortening delay thanks to less intermediate nodes between the origin server and the nodes. However, broadband of the streaming media server will be occupied by the massive AN. More importantly, occupying cloud service has to be paid, including transmission from the origin server to AN and from AN to the root node.
- Layer mode: All the AN form a tree structure with only one AN at the root. Data is pushed to the AN at the root by the media source, before it is pushed to each node at different layers. The layer mode makes use of advantages of cloud services, such as extremely high speed of gigabit fiber-optic used in node communication within cloud, and free internal communication.

V. KEY PROBLEMS OF P2P STREAMING MEDIA SYSTEM IN CLOUD ENVIRONMENT

The huge advantages of P2P are retained when constructing a mixed streaming media system from the perspective of network environment framework. But it will be a heavy cost to realize streaming media system purely by means of cloud computing, and bottlenecks in network transmission and broadband will be difficult to overcome. Major problems of P2P streaming media transmission system in cloud computing environment lie in broadband, streaming media data processing and QoS.

A. Broadband

Broadband not only has an immediate impact on media transmission, but also affects the choice of duplicate strategies within streaming media system. There is no large differences between different types of strategies when network broadband is sufficient; however, different duplicate strategies can result in performance discrepancy when network broadband is limited. High broadband environment is the fundamental prerequisite for guaranteeing duplicating quality as a result of high consumption of broadband during streaming media content duplicating.

The combination of cloud computing and P2P framework is an effective choice in resolving problem related to broadband. A mixed CloudP2P framework of streaming media application is put forward, which is composed of a cloud component and many peer-to-peer groups. The role of cloud lies in providing seeds and assisting terminal users in finding nodes with the same interest. Data can be exchanged among nodes in peer-to-peer groups. Compared with approaches purely based on cloud, the combination of cloud computing and P2P reduces costs of infrastructure and network broadband.

B. Streaming Media Data Processing

Streaming media data processing, especially coding and compressing, requires a large amount of computing

resources, while distributed parallel computing in cloud indicates direction in this regard. The input media content is distributed to processing nodes by means of cloud computing, before the processed content is merged. The use of distributed video compressing process shortens the time for coding and compressing, which is conducive to high-definition video.

Currently, main cloud computing data processing strategies include MapReduce and Dryad. For MapReduce strategy, a large task is divided into multiple small tasks to be treated with distributed processing, with the intermediate results retained in the local processing server. By merging the intermediate results, the final result is obtained, presents as linear processing mode on the whole. Bottleneck problems exist in dividing specific tasks and iterative algorithm despite infinite expansion of computing capacity by increasing computing nodes. For Dryad strategy, a parallel model based on Directed Acyclic Graph (DAG) is adopted. Each data processing task is represented by DAG, and the expansion of the processing method to two-dimensional mode makes data processing more flexible. In terms of streaming media system, despite some bottleneck problems in data processing, MapReduce has advantages in media processing, include simple blocking of media, coding and decoding the blocked media by increasing massive service computing nodes before combining them as a whole.

C. Quality of Services (QoS)

During the process of producing streaming media and transmitting it to users, the system needs to provide services as required by users, and meets their demand such as server availability, network performance, service corresponding time and real-time performance. Cloud computing platform provides a large number of users with all kinds of resource computing services, and the system of the platform is constructed with the guarantee from the service layer, which ensures that services bought by users can meet their demands.

QoS is always inversely proportional to costs. For streaming media system, QoS means high-performance play fluency and low play delay. Cloud computing is adopted in cloud-assisted P2P streaming media system to improve overall broadband to ensure the lowest limit of QoS. Adding two forms of nodes i.e. active nodes and passive nodes can accelerate content multiplying and improve the available upload broadband of the whole system.

AN and PN service costs can be calculated as follows:

Cost for one round of transmission (Trround) is defined as Can and the calculating formula is as follows:

$$C_{an} = C_{vm} + m \cdot C_{chunk}$$

In the formula, C_{vm} is the cost for one AN operating virtual machine during one round of transmission, C_{chunk} is the cost for one data block transmitted from AN to one user node (root node), and m is the quantity of data block uploaded by AN at one time. Uploading capacity of AN should be made full use of; hence $m=Up(an)$. Similarly, PN

cost C_{pn} during one round of transmission can be calculated as follows:

$$C_{pn} = C_{storage} + r \cdot (C_{chunk} + C_{req})$$

In the formula, $C_{storage}$ is storage cost, C_{req} is cost for getting data block from PN at one time, and r is the quantity of data block transmitted by PN at one time. C_{chunk} is the cost for one data block transmitted from PN to one user node. $C_{storage}$ can be omitted since only several-minute video streaming can be stored in PN and cloud storage is free.

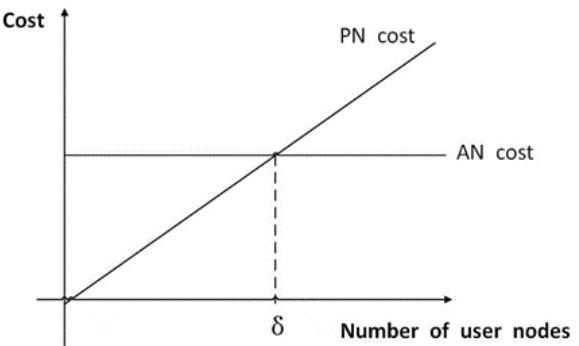


Figure 6. Relation between PN, AN and the number of user nodes

Relation between C_{an} , C_{pn} and the number of user nodes within one round of data block transmission (Trround) is shown in Figure 6. It can be observed that C_{pn} increases linearly as the number of user nodes increases. C_{an} is a constant regardless of the size of the cluster. Therefore, if intersection point δ of C_{an} and C_{pn} is decided, the reasonable time to increase AN instead of increasing PN load will be clear. δ is the balancing point of AN and PN costs.

$$\delta \approx \frac{C_{vm} + m \cdot C_{chunk}}{C_{chunk} + C_{req}}$$

VI. SUMMARY

The thesis studies P2P streaming media system network framework in clouding computing environment regarding to problems including limited bandwidth for P2P streaming media, insufficient computing resources and QoS. It summarizes integration of cloud computing and P2P and relevant technology, and indicates that existing bottleneck problems in network bandwidth can be overcome, high-efficiency transmission of streaming files can be achieved, play quality of audio and video files can be improved and stability of video resources can be guaranteed by P2P streaming media system based on cloud service.

REFERENCES

- [1] Peng Rui, Study of P2P Video Streaming Transmission Framework Combined with Cloud Service, 2014.
- [2] Yang Ge, Tian Hui, Key Technology of P2P Streaming Media Transmission in Cloud Environment, Computer Systems & Applications(2013), Vol 22, Issue 12.
- [3] Shen Jingbo, Study of Dynamic Content Management Based on P2P and Cloud Computing, 2011.

- [4] Wang Jundai, Design and Realization of P2P Streaming Media Platform Based on OpenStack, 2014.
- [5] Xu Jun, Study of P2P Video-on-Demand Technology in Cloud Environment, 2012.
- [6] Zhou Gemei, Liu Zebin, Application of P2P Streaming Media Technology Based on Cloud Service in Distance Teaching Video Transmission, Chinese Education Information-Elementary Education, 2016, Issue 3.