

Study on Distance Learning Scheme Based upon Peer to Peer(P2P)

Cuibo Yu*, Xuerong Gou*, Yong Wang**

**School of Network Education of Beijing University of Posts and Telecommunications, Room 512, DRC Mansion, No.28 Xijiekou Outer Street, Xicheng District, Beijing, P. R. China*

*** Software Development Center, Agricultural bank of China*

Jade_yu@sohu.com, xrgou@126.com, dixiechicks@sina.com

Abstract—Lifelong education is a tendency with the fast progress of society and technologies, and traditional educational forms including correspondence have some limits to meet the new requirements arisen from the tendency. P2P is one of contents distributed ways, in which all the hosts are treated equally and can access and make use of each other's resources such as storage of hard disk and CPU. The idea of combined P2P technology with e-learning platforms was brought forward to resolve the problem arisen from the lifelong education, which would decrease the time delay including retrieving and transmission, to cut down the bandwidth needed in this paper. The system structure was put forward and numerical analysis method was used to show the effect of the idea.

Keyword— P2P, peer to peer, e-learning, distance education, lifelong education.

I. INTRODUCTION

Lifelong education is a tendency with the fast progress of society and technologies, and traditional educational forms including correspondence have some limits to meet the new requirements arisen from the tendency. E-learning is the answer by combining the internet technology and education.

Modern distance learning (e-learning) is an educational form utilizing audio, video (live or VCR) and other computer technologies to send the course information out the campuses. It is one of new type of educational ways with the development of modern information technology (IT)[1], which integrates computer, multimedia and communication technologies. When talking about the concept of modern distance learning, the internet technologies and supporting computing platforms can't be avoided. Generally, traditional modern distance learning system adopts the C/S (Client/Server) or B/S (Browser/Server) architecture, as showed in figure 1.

Whether adopting C/S or B/S structure, there was/were only one or several main centralised server(s) in the system which clustered in the equipment rooms, and all the traffic of the communication messages and content distribution were dealt and carried by the centralised servers. At the beginning of the development of modern distance learning, the number of the registered students is not very large and the main servers in such a structure had the ability to process all the interactive flows. By now, much more people than ever need or feel the need to continue study to keep up with the fast

step of technologies, for all kind of reasons, some of them who can't leave the work to attend the campus class to study want to get the academic certificates, and some of them who graduated from the universities want to obtain more technique from online training, as thus the concept of lifelong education was brought forward. Under such situations, the number of people want to receive knowledge distributed through internet is very huge, the centralised C/S or B/S structure can't meet requirement of high processing capacity, and the main servers can't provide high throughput capability. It's time to modify the traditional centralised servers' structure to fit the new requirement with the social and economical development.

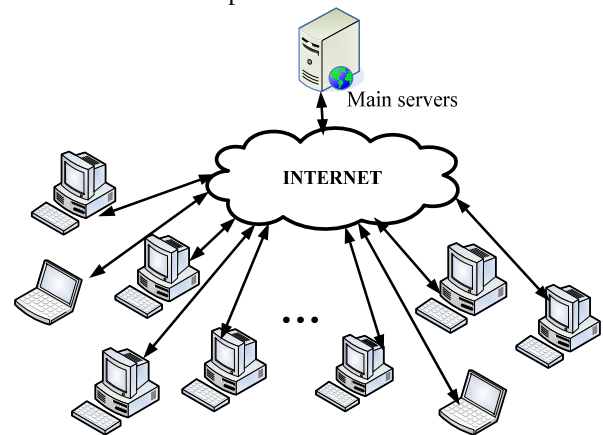


Figure 1 architecture of the generic distance learning system

Now the idea of Peer to peer (P2P) has been used in many fields such as file sharing, cooperative processing, information sharing and communication etc. According to the definition, the Peer-to-peer (P2P) network is a type of Internet network allowing a group of computer users with the same networking program to connect with each other for the purposes of directly accessing files from one another's hard drives, which made the idea of P2P can be adopted in the distance learning platform. The native quality that all the nodes are equal and distributed along the P2P overlay imitates the scenarios of the structure and essence of distance learning. There are three P2P topologies: unstructured, structured and hybrid network. No matter in which situation,

there are some peers which capability including bandwidth, uptime, CPU etc, are more powerful than the others. Such peers can chosen to be supernode (SN) [2,3,9,10]and were treated as sub-severs by the other peers, which provide contents (audio or video courseware) distribution for designated district. The combination of e-learning and P2P would be one of the means to resolve the problems resulted in lifelong education.

In this paper, the idea of layered P2P was used in the supporting system structure of distance learning. The paper was organized as following: the 1st part was the introduction of the research background, 2nd part was about the scheme of using P2P to set up a distance learning system, the performance of distance learning system based upon P2P was analyzed in 3rd part, the last part, i.e. the 4th part, was the conclusion of the whole paper and future work.

II. STRUCTURE OF THE SCHEME

There are two types of activities available in distance education systems, which are real-time and non-real-time separately. About real-time activities, the students are asked to “present” on-line at the designated time, which is requiring a timetable to be organized before. Because that all the participants in real-time system have to be present at the same time, generally, the clustered servers were used to concentrated control all the activities required by real-time, which was showed in figure 1.

As for non-real-time activities, such as BBS, courseware downloading etc, participants can access the materials on their own schedule. It’s not necessary to ask all students to be together at the same time. Under such circumstance, the distributed methods could be used to increase the accessing speed and decrease the processing delay.

P2P technology was one of the distributed ways. There are three structures in organized all the nodes in P2P overlay, in which only hierarchical P2P structure is suitable to be used in distance learning system. Layered P2P was one of the hierarchical structures. According to the layered P2P, the new scheme of the platform for content distribution in distance learning was showed in figure2.

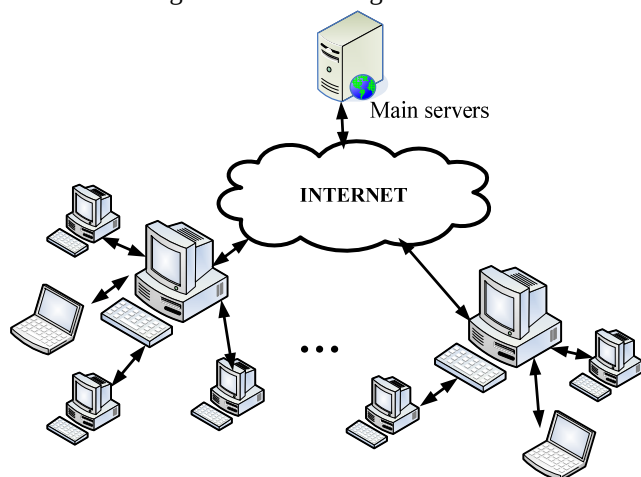


Figure 2 the scheme of distance learning based upon P2P

In the figure 2, all the clients accessing the server construct a P2P overlay, and on the basis of the capability of the peers and the gross area, the peers with relative powerful bandwidth and CPU would be elected as SN to be the sub-server of its district [2, 3]. After that, the SNs obtained the contents such as courseware and other related information from the main servers, and as long as the SNs are on-line, other peers around the SNs could retrieve these contents from the SNs. Therefore, the message burden among the whole network was reduced.

III. NUMERICAL ANALYSIS

Assume that there were 10,000 students registered in a network school, which providing distance educational service to the registered students. The number of 10,000 registered students is not very great for a normal network college. But it’s a great challenge for distance education supporting platform. Under the distance educational environment, the students obtained almost all the learning resources from the supporting platform (i.e. servers) by accessing internet. There are some basic learning activities, including downloading courseware and uploading homework. Such downloading and submitting operation requires the processing capacity of servers and the enough bandwidth along the transmitting path. We made assumption on the numerical value about the above mentioned activities as Table 1

Table 1. RelativeData of normal nodes

Item	Size of homework (MB)	Size of course-ware(MB)	Concurrence ratio of downloading	Concurrence ratio of uploading
Value	0.5	100	0.2	0.8

In traditional C/S or B/S structure as in Figure 1, the processing capacity of the main servers should be:

- Traffic of Downloading:
 $10,000 * 100 * 0.2 = 200,000MB$ (1)
- Traffic of Uploading:
 $10,000 * 0.5 * 0.8 = 4,000MB$ (2)

In table 1 and eqn. (1, 2), only the two familiar learning activities were considered, there came to such a huge traffic throughout the internet! If video/audio courseware and real-time interactive information was considered, it would be a much more thumping result. There needs very high performance servers and transmitting lines and other related equipments to process and transmit such huge traffic. Because all the servers are clustered around the same place, and all the students have to access the servers from all around the world to the same destination, any errors in the main servers or transmitting lines or would result in great problems. Moreover, heavy traffic transmission would induce long processing delay, which would worse the students’ e-study experience online. And bad learning

experience might make some students' give up the idea to acquire knowledge from distance educational system.

On the other side, if the layered P2P with SNs structure was adopted, the whole P2P overlay would be divided into several small districts according to the system's load, such as in the same case as Table 1, the gross registered students were divided into 20 small sub-overlays on the basis of geographical information or other rules. In each sub-overlay, one SN was elected according to its capability [2, 3]. Because the SNs were chosen as the agents to access the main servers, compared to the normal node, the concurrence ratio of SN would increase. Related data of layered P2P was assumed as table 2.

Table 2. RelativeData of SNs

Item	Size of homework (MB)	Size of course-work (MB)	Concurrence ratio of downloading of SNs	Concurrence ratio of uploading of SNs
Value	0.5	100	0.4	0.9

About downloading, only the SNs retrieved the needed resources from the main servers, other hosts (peers or nodes) in each sub-overlay got those contents from the chosen SN. Under such assumptions, the downloading traffic throughout the internet was showed as eqn. (3):

$$20 * 100 * 0.4 = 800MB \quad (3)$$

As eqn. (2) showed, only 0.2% of downloading traffic as in eqn. (1) throughout the Internet, and other information and interactive message flows were restrained only in the sub-overlays individually, which would increase the retrieving speed and decrease the time delay.

As for uploading, the normal nodes could upload directly to main servers, they also can upload to SN, and through SN upload the homework to main servers. Assume half nodes upload the content themselves and the other half through SNs. The traffic of internets looked as eqn. (4)

$$20 * 0.5 * 0.8 * 500 * 0.5 * 0.9 + 5000 * 0.5 * 0.8 = 3800MB \quad (4)$$

As eqn. (4) showed, only 200MB of uploading traffic as in eqn. (1) throughout the Internet was reduced by layered P2P structure, i.e., comparing with downloading, layered P2P gave minor improvement in decreasing the data flow of internet, its potential merits maybe cut down the transmitting time. Such a result was resulted from the following fact:

- The limited capability of SN: The uploading operation needed to write into the storage, and normal personal computer, even which have high configuration, maybe can't bear the high frequency writing operation. Therefore, the SN just helped those whose capability was very low to upload their contents online, here, 50% of the nodes in each sub-overlay were such class of peers.

- The security requirement of some students: Generally speaking, to students, the uploading material was used to examine their studying effect, which was more important than the downloading material. Therefore, they would try their best to guarantee the security. If they could find ways to access the internet, they preferred to do the uploading work themselves.

IV. CONCLUSION AND FUTUREWORK

In this paper, the idea of layered P2P technology was adopted to resolve the problem of high bandwidth and processing capacity to database servers arisen from the great number of users in the computing platform of lifelong education, which would decrease the time delay including retrieving and transmission, to cut down the bandwidth needed. Limited by the development of layered P2P technology, only simple numerical analysis was used to illuminate the effect of idea in this paper. In the future, more work should be done to analyse the merits of layered P2P technology to distance education.

Because lifelong education would be the long-term tendency with the fast progress of society and technology, there needs to do more research on related implementation technologies to accelerate the permeation and popularization of lifelong learning. Except P2P, there are more distributive ways needed to investigate if they can be used in lifelong education.

ACKNOWLEDGMENT

The authors would like to thank the colleagues of School of Network Education for their valuable discussions and suggestions, and give greatly thanks to all the anonymous reviewers.

This material is based upon project supported by Chinese Universities Scientific Fund 2009RC0902.

REFERENCES

- [1] Distance education, http://en.wikipedia.org/wiki/Distance_education
- [2] G. Wang; Cb Yu; et al Supernode Election Algorithm in DHT Based upon Network Coordinate and District Partition, International Conference on Communication Software and Networks, 2009. ICCSN '09. ,27-28 Feb. 2009 Page(s):316 – 321
- [3] Cb Yu; Xr Gou; et al; Study on Supernode Election Algorithm in P2P Network Based upon District Partitioning, International Conference on Communication Software and Networks, 2009. ICCSN '09. 27-28 Feb. 2009 Page(s):196 - 199
- [4] V. Lo, D.Y. Zhou et al, Scalable Supernode Selection in Peer-to-peer overlay Networks, Hot Topics in Peer-to-peer Systems, Second International Workshop on 21 July 2005, PP.18-25.
- [5] D. Stutzbach, R. Rejaie, Understanding Churn in Peer-to-Peer Networks, Internet Measurement Conference Proceedings of the 6th

- ACM SIGCOMM conference on Internet measurement Rio de Janeiro, Brazil SESSION: Peer to peer, Pages: 189 - 202
- [6] J.w. Shi, Y. Wang, et al, A Hierarchical Peer-to-Peer SIP System for Heterogeneous Overlays Interworking Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE, page(s): 93-97
- [7] S. Singh and J. Kureose, Electing Leaders Based Upon Performance: the Delay Model, Distributed Computing Systems, 1991., 11th International Conference on Volume , Issue , Page(s):464 - 471
- [8] A. Singla, C. Rohrs, Ultrapeers: Another Step Towards Gnutella Scalability, Lime Wire LLC, Working Draft, <http://www.peer-to-peer.info/bibliography/singla2002ultrapeers>
- [9] Y.M. Liu, S.B. Yang et al, the research of the reputation-aware super node selection algorithm in P2p system Journal of the graduate school of the Chinese academy of sciences, 2008.3
- [10] I. Stoica, R. Morris, et al, Chord: A scalable peer-to-peer lookup service for internet applications in SIGCOMM, San Diego, CA, USA, Aug 2001.
- [11] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in P2P networks .In Proc. of the Twelfth International World Wide Web Conference (WWW2003). 2003.
- [12] Selcuk A A,Uzun E,Pariente M R.A reputation-based trust management system for P2P networks.Cluster Computing and the Grid,CCGrid 2004[C] IEEE International Symposium on,2004,251-258.